EPFL

Thèse n° 8042

The molecular treasure box of environmental *Vibrio cholerae* strains

Présentée le 1er septembre 2021

Faculté des sciences de la vie
Unité de la Prof. Blokesch
Programme doctoral en approches moléculaires du vivant

pour l'obtention du grade de Docteur ès Sciences

par

Natália Carolina DREBES DÖRR

Acceptée sur proposition du jury

Prof. B. Lemaitre, président du jury
Prof. M. Blokesch, directrice de thèse
Prof. F. Le Roux, rapporteuse
Prof. D. Salomon, rapporteur
Prof. M. Pilhofer, rapporteur

École
polytechnique
fédérale
de Lausanne

2021

“Somewhere, something incredible is waiting to be known.”

Carl Sagan

To my love, Matheus;

my wonderful parents, Walter and Viviane;

and my other half in the world and sister, Julia.

ABSTRACT

Cholera pandemics have been affecting humankind for centuries and are still considered a major public health problem, especially in regions around the world with poor access to clean water and sanitation. Cholera pandemics are caused by a specific lineage of the bacterial pathogen *Vibrio cholerae*, while the vast majority of the species' diversity is found in innocuous environmental strains. The complete set of factors that allowed the emergence and success of the pandemic lineage are still unknown. Importantly, most *V. cholerae* strains are adapted to aquatic habitats and trigger their behavior accordingly. The aim of this thesis was therefore to gain a better understanding of what is special about the pandemic lineage when contrasted to its environmental counterparts. Specifically, we investigated environmentally important bacterial behaviors in a comparative framework. Firstly, considering the relevance of eukaryotic predation in shaping bacterioplankton structure, we investigated possible strategies used by *V. cholerae* to avoid grazing by the predatory amoeba *Dictyostelium discoideum*. We observed that *V. cholerae* was able to rapidly intoxicate amoebae and that the intoxication dynamics differed when caused by pandemic or environmental strains. Secondly, we compared a well-conserved set of mobile genetic elements (MGEs) of the pandemic lineage, many of which are involved in key aspects of cholera pathogenesis, with the respective counterparts of environmental strains. To do so, we sequenced and *de novo* assembled the genomes of fifteen environmental isolates and thoroughly described their mobilome. Moreover, transcription profiling demonstrated that most genes located on these MGEs are expressed. Thirdly, we compared environmental and pandemic strains regarding their potential to kill bacterial competitors and to defend themselves against protozoan predation. Specifically, we assessed the role of two molecular weapons: the pore-forming toxin hemolysin and the type six secretion system (T6SS), a molecular killing device that delivers effector toxins into target cells. While environmental

strains keep both of these weapons constitutively active, pandemic strains employ regulatory mechanisms to control their expression. We observed that all environmental *V. cholerae* isolates used their T6SS to efficiently outcompete prey bacteria, while only two clades of the environmental isolates also intoxicated eukaryotic amoebae. Furthermore, we performed a meticulous *in silico* characterization of the effector and immunity proteins carried by the environmental strains and showed in pairwise killing experiments that a high degree of immunity protein identity was required to allow the strains' coexistence. Finally, we addressed the phenomenon of T6SS constitutive activity in non-pandemic *V. cholerae* strains, which has puzzled the field since the T6SS's discovery in 2006. Using a transformation-based strain library, we uncovered a single nucleotide polymorphism that controls T6SS activity. In summary, work developed in this thesis contributes to a better understanding of *V. cholerae*'s evolution from an innocuous inhabitant of aquatic environments to a pandemic-causing human pathogen.

Keywords: Pathogen emergence, *Vibrio cholerae*, amoebae, host-pathogen interactions, mobile genetic elements, type VI secretion system, regulatory networks

RESUMÉ

Les pandémies de choléra affectent le genre humain depuis des siècles et sont toujours considérées comme un problème de santé publique majeur, en particulier dans les régions du monde où l'accès à l'eau potable et aux systèmes d'assainissement est limité. Les pandémies de choléra sont causées par une lignée spécifique du pathogène bactérien *Vibrio cholerae*, tandis que la grande majorité de la diversité de cette espèce est constituée des souches environnementales inoffensives. L'ensemble complet des facteurs qui ont permis l'émergence et le succès de la lignée pandémique est encore inconnu. La plupart des souches de *V. cholerae* sont adaptées aux habitats aquatiques et déclenchent leur comportement en conséquence. Le but de cette thèse était donc d'acquérir une meilleure compréhension des caractéristiques spécifiques de la lignée pandémique par rapport à ses homologues environnementaux. Plus précisément, nous avons comparé les comportements bactériens importants pour la survie dans l'environnement chez des souches pandémiques et environnementales de *V. cholerae*. Tout d'abord, considérant l'importance de la prédation eucaryote sur la composition du bactérioplancton, nous avons étudié les stratégies possibles utilisées par *V. cholerae* pour résister à l'amibe prédatrice *Dictyostelium discoideum*. Nous avons observé que *V. cholerae* était capable d'intoxiquer rapidement les amibes et que la dynamique d'intoxication différait lorsqu'elle était causée par des souches pandémiques ou environnementales. Deuxièmement, nous avons comparé un ensemble bien conservé d'éléments génétiques mobiles (MGEs) de la lignée pandémique, dont beaucoup sont impliqués dans des aspects clés de la pathogenèse du choléra, avec les homologues respectifs des souches environnementales. Pour ce faire, nous avons séquencé et assemblé *de novo* les génomes de quinze isolats environnementaux et décrit minutieusement leurs MGEs. De plus, nous avons montré que la plupart des gènes situés sur ces MGEs sont exprimés au niveau transcriptionnel.

Troisièmement, nous avons comparé les souches environnementales et pandémiques quant à leur potentiel à tuer les concurrents bactériens et à se défendre contre la prédation par les protozoaires. Plus précisément, nous avons évalué le rôle de deux armes moléculaires : l'hémolysine, une toxine formant des pores et le système de sécrétion de type six (T6SS), un dispositif moléculaire qui délivre des effecteurs toxiques dans les cellules cibles. Alors que les souches environnementales maintiennent ces deux armes constitutivement actives, les souches pandémiques utilisent des mécanismes de régulation pour contrôler leur expression. Nous avons observé que tous les isolats environnementaux de *V. cholerae* utilisaient leur T6SS pour supplanter efficacement les bactéries cibles, tandis que seulement deux clades des isolats environnementaux intoxiquaient également les amibes eucaryotes. En outre, nous avons effectué une caractérisation méticuleuse *in silico* des protéines effectrices et immunitaires portées par les souches environnementales et montré dans des tests de prédation inter-bactériens qu'un degré élevé d'identité des protéines immunitaires entre deux souches était nécessaire pour permettre leur coexistence. Enfin, nous avons étudié le phénomène de l'activité constitutive du T6SS dans les souches non pandémiques de *V. cholerae*, qui a intrigué le domaine depuis la découverte du T6SS en 2006. En utilisant une banque de souches obtenues par transformation, nous avons une région génomique unique qui contrôle l'activité du T6SS. En résumé, les travaux développés dans cette thèse contribuent à une meilleure compréhension de l'évolution de *V. cholerae* d'un habitant inoffensif des milieux aquatiques à un pathogène humain à l'origine d'une pandémie.

Mots clés : émergence d'agents pathogènes, *Vibrio cholerae*, amibes, interactions hôte-pathogène, éléments génétiques mobiles, système de sécrétion de type VI, réseaux de régulation

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF SUPPLEMENTARY FIGURES

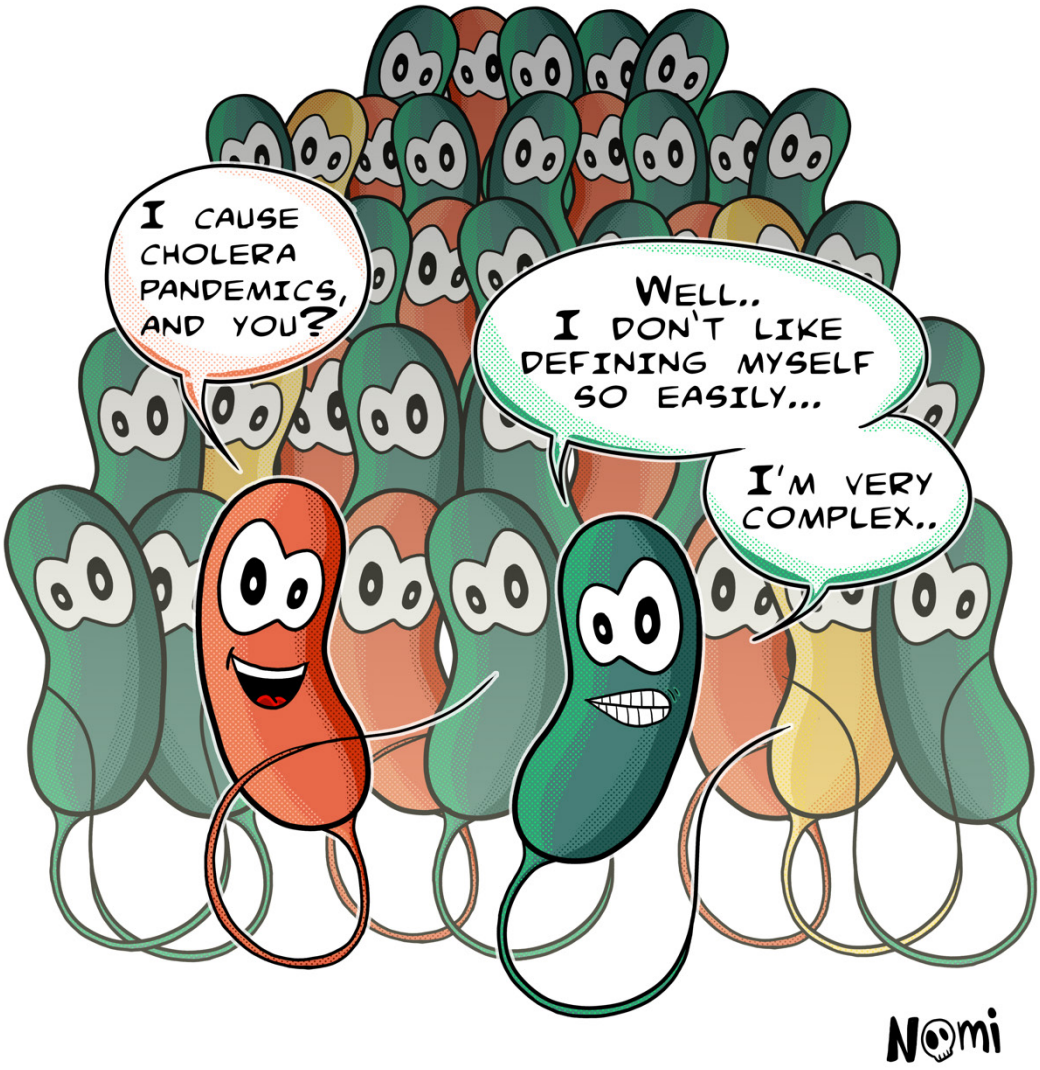
LIST OF SUPPLEMENTARY TABLES

LIST OF ABBREVIATIONS

ACD	Actin-Crosslinking Domain
Amp	Ampicillin
aph	Aminoglycoside phosphotransferase gene (kanamycin resistance)
Aux	T6SS Auxiliary cluster
BCCT	Betaine/ Choline/ Carnitine Transporter
bp	Base pairs
cat	Chloramphenicol acetyltransferase (chloramphenicol resistance)
Cat	Chloramphenicol
cAMP	Cyclic adenosine monophosphate
CDI	Contact-Dependent growth Inhibition
c-di-GMP	Cyclic diguanylate monophosphate
CFTR	Cystic Fibrosis Transmembrane conductance Regulator
CFU	Colony Forming Units
cGas	Cyclic GMP-AMP synthase
chr	Chromosome
COVID-19	Coronavirus disease
CRISPR	Clustered Regularly Interspaced Short Palindromic Repeats
CTX	Cholera toxin
CTXΦ	Cholera toxin bacteriophage
CV	Contractile Vacuole
DASW	Defined Artificial Seawater
EFV	Expelled Food Vacuole
E/I	Effector/ Immunity protein pair
EPS	Exopolysaccharide
ER	Endoplasmic Reticulum
FRT	Flippase Recognition Target
gDNA	Genomic DNA
Gent	Gentamycin
HCD	High Cell Density
HGT	Horizontal Gene Transfer
IS	Insertion Sequence
kb	Kilo base pairs
kDa	Kilo Dalton
MGE	Mobile Genetic Element
MSHA	Mannose-Sensitive Hemagglutinin Pilus
nan-nag	Sialic acid utilization cluster

OD_{600}	Optical Density measured at 600nm
ORF	Open Reading Frame
PAI	Pathogenicity Island
PBS	Phosphate Buffered Saline
PCR	Polymerase Chain Reaction
PGAP	Prokaryotic Genome Annotation Pipeline
PTLB	Phage-Tail Like Bacteriocins
QS	Quorum Sensing
R/M	Restriction/ Modification system
SD	Standard Deviation
SNP	Single Nucleotide Polymorphism
STING	Stimulator of Interferon Genes
Strep	Streptomycin
T2SS	Type II Secretion System
T4P	Type IV Pilus
T4SS	Type IV Secretion System
T6SS	Type VI Secretion System
Tad pilus	Tight adherence pilus
TCP	Toxin Coregulated Pilus
Tn	Transposon
VPI-1	*Vibrio* Pathogenicity Island 1
VPI-2	*Vibrio* Pathogenicity Island 2
VSP-I	*Vibrio* Seventh Pandemic island I
VSP-II	*Vibrio* Seventh Pandemic island I
WGS	Whole Genome Sequencing
WHO	World Health Organization
WT	Wildtype

1 GENERAL INTRODUCTION



Drawing by Noémie Matthey, PhD

1.1 PATHOGEN EVOLUTION IN THE ENVIRONMENT

The burden exerted by infectious diseases on human kind has probably never been clearer for our generation than at the present moment, while the whole world experiences the effects of a coronavirus disease (COVID-19) pandemic caused by the SARS-CoV-2 virus. A major point of concern in the study of infectious diseases, especially those that cause pandemics, is to understand how these microorganisms evolved from their environmental progenitors (1, 2). Even though direct adaptive selection that increases virulence is pivotal in this sense, it would be presumptuous to assume bacteria evolved intricate virulence mechanisms to deal specifically and uniquely with humans as their host. In this context, the coincidental evolution hypothesis, initially proposed by Levin and Svanborg Edén in 1990 (3), suggests that many virulence factors were originally selected in an environmental context and not the final host where disease develops, which, ultimately, has been experimentally demonstrated for a subset of pathogens (4–7). In the environment, bacteria have been fighting against many burdens over an evolutionary speaking very long time. These burdens include inadequate abiotic conditions, competition with other microorganisms, and predation by bacteriophages and protozoans, to name a few (8). Some of these adaptations were potentially coopted to operate as virulence factors in a human disease context.

1.2 *VIBRIO CHOLERAE* AS A MODEL FOR LIFESTYLE SHIFT

Vibrio cholerae, a comma-shaped Gram-negative bacterium, is part of the Vibrionaceae family, which belongs to the class of Gammaproteobacteria. The Vibrionaceae family includes 12 genera with nearly 140 species, and all members have the distinctive feature of containing two chromosomes (9). The genus *Vibrio* is abundant in marine environments and species can be found associated to a wide diversity of organisms, such as sponges, corals, phyto- and zooplankton, fish, shellfish, etc. While some species are symbionts, many of

them are important pathogens for marine organisms or humans, causing vibriosis and gastrointestinal diseases, respectively (9, 10).

Among these species, *V. cholerae* is a great model to study pathogen emergence. While being famously recognized as the causative agent of deadly disease cholera (11), *V. cholerae* is actually a diverse species that primarily inhabits the aquatic environment. In these locations, *V. cholerae* can be found as planktonic free-living bacteria or as biofilm-associated bacteria attached to chitinous surfaces from zooplankton, such as the (molted) exoskeletons of small crustaceans (i.e., copepod; (Figure 1.1, left)). Cholera-causing pathogenic strains of *V. cholerae* display adaptations to colonize the human gastrointestinal tract (Figure 1.1, right). However, the vast majority of *V. cholerae* diversity is composed of environmental isolates. These isolates do not harbor the cholera-causing virulence factors, though they can be sparsely associated to cases of gastrointestinal illness (12–18).

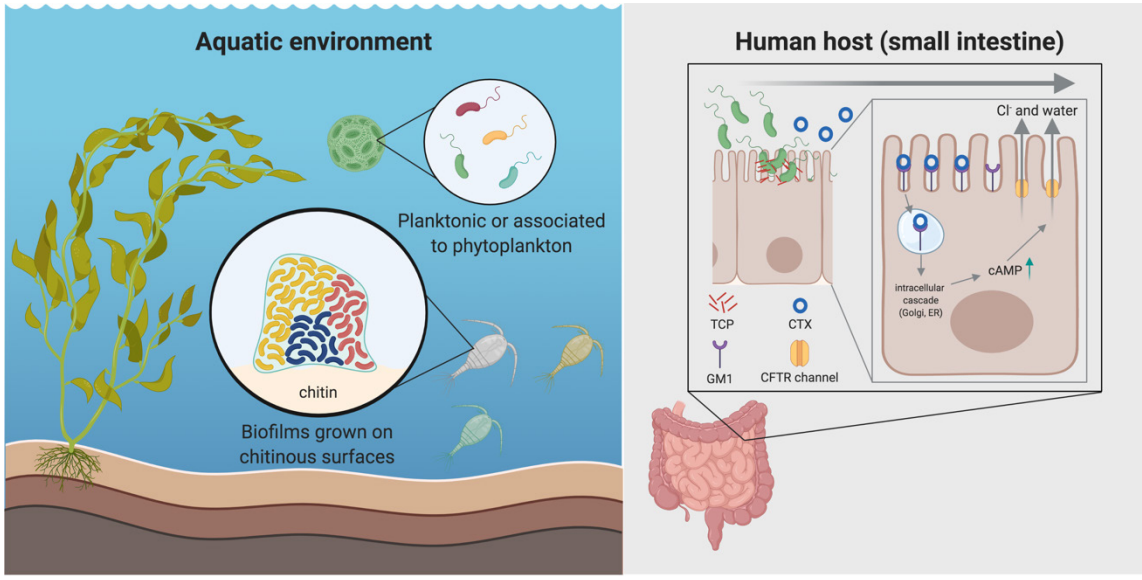


Figure 1.1 *V. cholerae* and its dual lifestyle.

V. cholerae bacteria are major inhabitants of the aquatic environment (left), where they can be found either as free-living bacteria or as biofilms growing on chitinous surfaces, such as the exoskeleton of small crustaceans. Conversely, toxigenic strains of *V. cholerae* are able to colonize the human small intestine (right). Production of the toxin-coregulated pilus (TCP) allows microcolony formation and adherence to the gut epithelium. Cholera toxin (CTX) is secreted in the intestinal milieu and binds to GM1 ganglioside receptors, leading to its internalization and processing, mostly in the endoplasmic reticulum (ER). A series of processes lead to CTX-dependent increased of cyclic AMP (cAMP) levels by the indirect activation of adenylate cyclase. Elevated

cAMP concentrations lead to efflux of chloride through the cystic fibrosis transmembrane conductance regulator (CFTR) channel and the concomitant efflux of copious amounts of water. Figure based on (17, 19, 20), created with BioRender.com.

1.2.1 Cholera disease

Cholera is an acute diarrheal disease transmitted by contaminated water and food. It is therefore of special concern in parts of the world with limited access to hygiene, sanitation, and clean drinking water. Cholera is an ancient disease that has been part of humanity for centuries, even though the first recorded cholera pandemic started only in 1817 (18). Since then, cholera has spread throughout the world in seven pandemics, the seventh of which started in 1961 and is currently ongoing (12, 17) with recent estimates of up to 4 million people who are infected every year (11).

Cholera pathogenesis develops with oral ingestion of *V. cholerae* from contaminated food or water, and this “inoculum” passes through a bottleneck at the low pH of the stomach. *V. cholerae* bacteria are quite sensitive to low pH, especially when free-living (20, 21). Volunteer studies in the 1970s demonstrated that the infection dose required to establish cholera-like disease was around 10^{11} bacteria, and this dose was lowered to 10^{6} when the gastric pH was neutralized (22). After passing through the stomach, the bacteria reach the epithelium of the small intestine by flagellar-based motility and use their toxin-coregulated pili (TCP), a primary colonization factor, to auto-aggregate and adhere to the intestinal epithelium thereby forming microcolonies (Figure 1.1, right) (23, 24). TCP is also the receptor for the cholera toxin (CTX) bacteriophage (CTXΦ), which carries the genes (*ctxAB*) that encode CTX (25). CTX is an AB_5-subunit toxin. A single A subunit, which carries the toxic activity, is attached to a ring of five B subunits, which are responsible for receptor binding. After its secretion by the pathogen, CTX binds to the GM1 gangliosides on epithelial cells followed by its transport inside the cell. Next, the toxin undergoes a trafficking cascade

through the endoplasmic reticulum-associated protein degradation (ERAD) pathway, after which it reaches the cell's cytosol where it indirectly fosters the increased production of cyclic AMP (cAMP). High cAMP concentration leads to higher levels of chloride and bicarbonate export (through the cystic fibrosis transmembrane conductance regulator – CFTR) and loss of copious amounts of water, which is the hallmark of cholera symptoms (17, 26).

1.2.2 Pandemic lineage emergence

The two main cholera virulence factors, TCP and CTX, are encoded on mobile genetic elements, namely the TCP island (or *Vibrio* Pathogenicity Island 1 – VPI-1) and the CTX prophage. Consequently, the ability of a *V. cholerae* strain to cause at least the most crucial steps in cholera pathogenesis depends on the acquisition of these elements, and strains that harbor them are referred to as "toxigenic" strains. Interestingly, however, not all toxigenic strains are able to cause cholera pandemics. In fact, all pandemic *V. cholerae* strains belong to only two out of the around 200 O-antigen serogroups known, namely O1 and O139. The O1 serogroup can be further classified into biotypes: classical and El Tor. O1 classical strains were responsible for the six previous cholera pandemics, while the O1 El Tor biotype is accountable for the 7th ongoing cholera pandemic (12, 14, 17). O139 strains are genetically derived from the 7th pandemic O1 El Tor strains (27), but are nowadays rarely associated to cholera outbreaks (12, 17).

Pandemic *V. cholerae* strains have very similar genomic content (28–31). Besides the TCP island and the CTX prophage, 7th pandemic strains contain a conserved set of mobile genetic elements (Figure 1.2). *Vibrio* Pathogenicity Island 2 (VPI-2) is a 57-kb element that contains a sialic acid utilization cluster (*nan-nag*), a type I Restriction-Modification (RM) system and a Zorya phage defense system (32, 33). The neuraminidase

encoded by a gene inside the *nan-nag* region has a role in cholera pathogenesis as it cleaves off the sialic acid from intestinal GM1 ganglioside receptors, allowing binding by CTX (33, 34). Furthermore, these strains harbor two *Vibrio* Seventh Pandemic islands (VSP-I and -II) (10, 35–37). VSP-I is 16 kb-long and contains a gene encoding a novel class of dinucleotide cyclase (DncV) (37). DncV was demonstrated to be an ancestral cGas protein, and the encoding *dncV* gene is part of a 4-gene antiphage defense system that is commonly found throughout the bacteria (38). Importantly, cGas/DncV was later shown to work in concert with a bacterial functional homolog of STING (39). The cGAS-STING pathway is ubiquitous in eukaryotes and has been extensively studied. Finally, VSP-II is 27 kb-long and, so far, has not been demonstrated to be a *bona fide* pathogenicity island, given that it contains mostly hypothetical genes (10, 40).

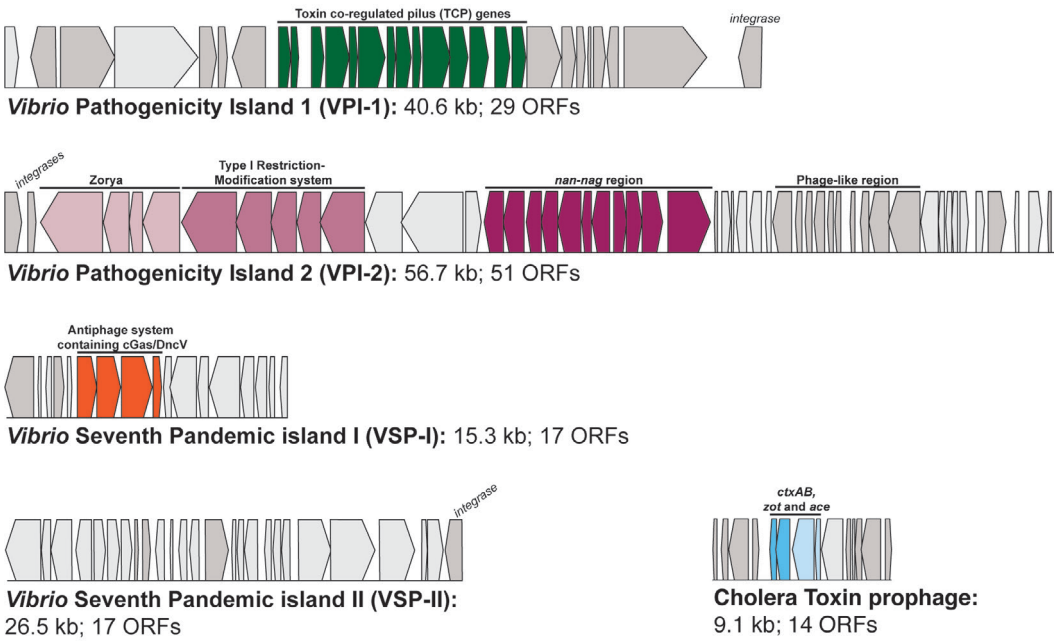


Figure 1.2 Main genomic islands in 7th pandemic *V. cholerae*.

7th pandemic strains of *V. cholerae* contain a conserved set of genomic islands. VPI-1 contains the *tcp* genes that encode the toxin co-regulated pilus (TCP), important for bacterial auto-aggregation and adherence to the intestine and is the receptor for cholera toxin (CTX). VPI-2 contains a Zorya phage defense system, a type I Restriction-Modification (RM) system, a phage-like region and a sialic acid utilization cluster (*nan-nag*). VSP-I and VSP-II contain mainly genes encoding hypothetical proteins. However, VSP-I contains a 4-gene cluster that includes a cGas/ DncV encoding gene. This protein is part of a pathway ancestral to the eukaryotic cGAS-STING antiviral pathway. The CTX prophage contains the genes (*ctxAB*) encoding CTX and the accessory

toxins Zot and Ace. Dark grey arrows represent genes that have known annotations but that are not labeled in this figure for clarity. Light grey arrows indicate genes encoding hypothetical proteins.

The astounding and understudied diversity found in environmental *V. cholerae* (16, 28, 41) when compared to pandemic isolates showcases how specific the pandemic lineage is. The acquisition of the TCP island by toxigenic isolates, allowing the lysogenic conversion by the CTXΦ, is a clear example of direct selection of virulence, since this acquisition was pivotal for *V. cholerae* to become a pathogen (5, 25, 42). However, that alone was not a deterministic factor for cholera pandemicity. In fact, there have been reports of toxigenic non-pandemic strains causing small and localized cholera outbreaks that never reached pandemic levels (14, 28, 43–45). Furthermore, water sampling throughout the world have revealed instances of environmental TCP- and/or CTX-positive isolates (46–51). The full suite of factors responsible for the emergence and success of the pandemic lineage is not understood and most likely involves elements such as outbreak location and the sanitation and social conditions encountered, the strain genomic content, and strain-specific phenotypes. Going back to the coincidental evolution hypothesis and considering that *V. cholerae* is primarily an aquatic bacterium, it is therefore plausible that environmentally-derived adaptations could have shaped some aspects of the pandemic lineage.

Bacteria like as *V. cholerae* face multiple selective pressures in the aquatic environment, such as inadequate access to nutrients, grazing pressure from predators and competition with other bacteria. Bacteria have developed an astonishing set of mechanisms to adapt to these pressures, which will be briefly discussed below.

1.3 PREDATOR EVASION STRATEGIES

Predation by bacteriophages and protozoans has been one of the major environmental selective forces experienced by bacteria, most likely since billions of years (8, 52). Lysis driven by phages and grazing by protists each account for about half of microbial mortality in the aquatic environment (8, 52–54). While protozoan predation plays a role mainly in limiting total abundance of bacterioplankton biomass, bacteriophages are more specific and therefore tend to affect community diversity (54). For the sake of the argument of this thesis, we shall now focus on the protozoan predators.

Bacteria have evolved a wide range of mechanisms to circumvent predation by protozoa (8). The amoebal feeding process typically involves forcing grazed bacteria into a food vacuole. It is therefore possible to classify bacterial evasion strategies as either occurring before ingestion (extracellularly) or from inside the food vacuole (intracellularly) (Figure 1.3). Simple yet effective extracellular adaptations include morphological plasticity, such as oversize and filamentation (8, 52, 53, 55) (Figure 1.3a), and increased motility (52, 53) (Figure 1.3b). Moreover, as most protozoa require specific receptors on the bacterial surface in order to start phagocytosis, bacteria can “hide” from predators by modifying these receptors (Figure 1.3c). Furthermore, they can use their lipopolysaccharide as a physical protection (8, 52, 53). Additionally, bacteria can secrete exopolymers and therefore become encased in microcolonies and biofilms, which can enhance bacterial survival against protozoan grazing (52, 56) (Figure 1.3d). Importantly, when interacting in biofilms, bacterial populations can synchronize their behavior through quorum sensing (QS). That ultimately results in a cooperative defense response, such as the release of toxins (Figure 1.3e), increasing fitness upon predation (8, 52). Moreover, some of these cooperative defenses with an ecological function might be coopted and have a role as virulence factors in a human disease context (8, 52, 57).

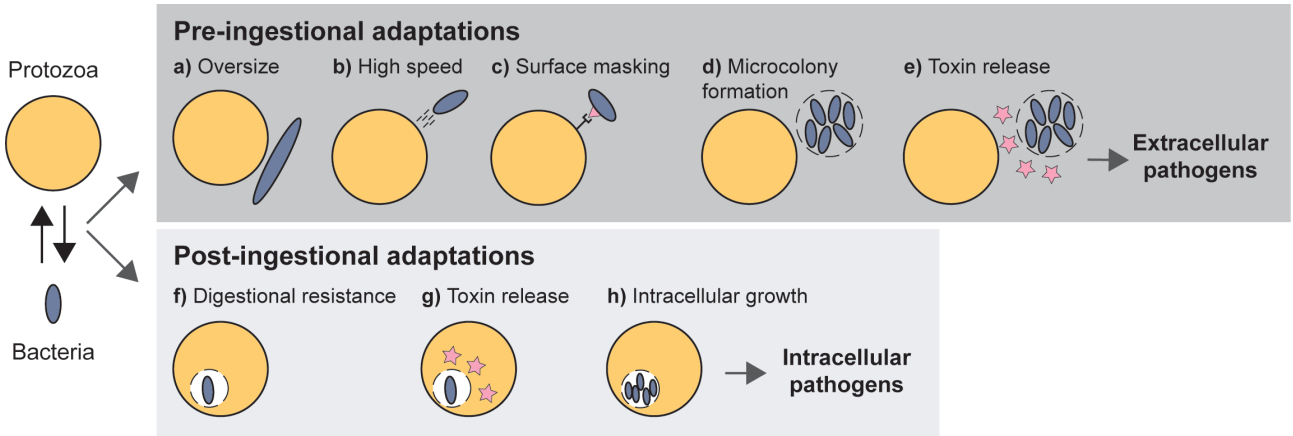


Figure 1.3 Bacterial adaptations against protozoan predation.

Anti-predation strategies emerging from bacteria-protozoa interactions. Upper rectangle showcases pre-ingestional adaptations (a-e), with increasing degrees of complexity. Some of these adaptations can function as virulence factors in a disease context (origin of extracellular pathogenesis). Lower rectangle depicts post-ingestional adaptations (f-h) employed by bacteria from inside a phagosome. These adaptations can lead to the evolution of intracellular pathogens. Freely adapted from (52).

Once ingested by protozoa, bacteria encounter harsh conditions in the phagosome (food vacuole), such as acidification, oxidative stress, nutrient deprivation, antimicrobial compounds and digestive enzymes (58). Bacteria can subvert these conditions by manipulating the phagosome-lysosome pathway, which, ultimately, allows them to use these predatory cells as a replication niche (Figure 1.3f-h). In that sense, protozoa can serve as environmental reservoirs and training grounds for the evolution of intracellular pathogens that can potentially also infect other types of phagocytic cells, such as macrophages (6, 8, 58–60). For example, *Legionella pneumophila*, *Brucella abortus,* and *Chlamydia* species can redirect the phagosome right after uptake, leading to the formation of a vacuole that does not have endocytic features. Other intracellular pathogens can block the phagosomal maturation pathway at later steps. For example, *Mycobacterium tuberculosis* can arrest the pathway at the pre-phagolysosome step, avoiding fusion with lysosomal contents. *Listeria monocytogenes*, *Shigella flexneri* and *Francisella tularensis* can break the maturing

phagosome and escape into the cytosol, which they can use for nutritious purposes (4, 6, 8, 52, 58, 60–66).

1.3.1 Anti-predation molecular weaponry of *V. cholerae*

There have been studies showcasing some of the strategies used by *V. cholerae* to resist predation by different protozoa. For example, Matz and collaborators (56) observed that biofilm-associated *V. cholerae* were resistant to feeding by the surface-grazing flagellate *Rhynchomonas nasuta*, while planktonic cells were almost completely consumed by the suspension-feeding flagellate *Cafeteria roenbergensis*. Their results also demonstrated increased biofilm formation in planktonic cells as a response against grazing pressure (56). A follow-up study indicated that chitin-grown *V. cholerae* biofilms were able to inhibit the growth of *R. nasuta.* This inhibition was explained by the QS-dependent production of ammonium, a by-product of chitin metabolism (67). Vaitkevicius *et al.* discovered that *V. cholerae* was able to cause lethal infections in *Caenorhabditis elegans* worms, and demonstrated that the QS-dependent production of a specific protease (PrtV) was pivotal for this phenotype. This protease was also shown to play a role in the protection against predation by the ciliate *Tetrahymena pyriformis* and the flagellate *C. roenbergensis* (68). Finally, pioneer work that lead to the discovery of the type VI secretion system (T6SS) showed that toxigenic non-pandemic strains of *V. cholerae* employ an actin cross-linking effector delivered by the T6SS to intoxicate *Dictyostelium discoideum* amoebae and macrophages (69–71).

Even though toxigenic *V. cholerae* strains are able to colonize the human small intestine, they are still primarily aquatic bacteria which can be found associated to diverse organisms in the environment (15, 16, 18). Importantly, considering that cholera is mainly

transmitted by contaminated water, potential environmental reservoirs are of major epidemiological concern, especially during in-between-outbreak periods (16, 18, 72). In that sense, a few studies have demonstrated the ability of *V. cholerae* to resist intracellular digestion by protozoa, potentially using these organisms as persistence hotspots in the environment. Our laboratory has demonstrated that *V. cholerae* is capable of resisting intracellular digestion by the aquatic amoeba *Acanthamoeba castellanii.* The ingested/phagocytosed bacteria are then released back into the environment. However, work in our group also showed that an alternative pathway exists in which food vacuoles fuse with an osmoregulatory organelle known as the contractile vacuole (CV), thereby establishing a replication niche. This niche lasts throughout amoebal encystation, allowing bacteria to actively replicate. During this time, proper regulation of the HapA protease, which is responsible for cleaving the pore-forming toxin hemolysin, is essential to avoid premature intoxication of the host. Finally, release of the enzyme lecithinase by *V. cholerae* allows lysis of the amoebal cyst, and flagellar motility is then used by the bacteria to escape the succumbed host (73, 74). It has also been demonstrated that *V. cholerae* can resist adverse conditions in the phagosomes of *T. pyriformis* by employing the outer membrane protein OmpU. In this case, the large amount of non-digested cells inside the phagosomes is thought to trigger the expulsion of bacteria-containing vacuoles from these ciliates, termed expelled food vacuoles (EFVs). Being encased in EFVs can confer survival advantages to *V. cholerae*, as EFVs work as a protection against adverse conditions found in the environment or, potentially, in the human gastrointestinal tract. Indeed, the authors demonstrated that EFVs outcompeted planktonic cells in a mouse colonization model, possibly improving survival when passing through the acid gastric pH or when exposed to antimicrobial defenses (75).

1.4 BACTERIAL ANTAGONISM STRATEGIES

Bacteria are extremely abundant on Earth, with richness estimations of around 1.2 x 10^{30} bacterial (and archaeal) cells (76). These organisms are primarily found entrenched into dense communities in the ocean and soil (76), where they compete (and cooperate) by dynamically responding to fluctuating environmental conditions and resource availability (77, 78). The nature of their possible competitive interactions can vary. Exploitative or scramble competition refers to the rapid use of available resources without direct interaction among competitors. In these situations, bacteria might put in place relatively "passive" strategies to increase their chance of reaching those resources, such as motility, selective attachment and production of exopolysaccharides (EPS) (79). On the other hand, contest or interference competition refers to the active fight among competitors for the resource. In fact, it is now evident that bacteria invest a lot of their genomic space and energy into an arsenal of molecular weapons used for that end (78–80).

Bacterial weaponry can be generally classified by how they inflict damage in their competitors. Mechanical weapons, such as tailocins (also called phage-tail like bacteriocins, PTLBs), are able to punch holes and thereby physically damage cells, leading to rapid cell death (81, 82). Importantly, the attacker cell needs to lyse itself in order to release such tailocin particles (81). Biological weapons refer to phages, which are viruses that infect bacteria. The prophage dormant form, which is found integrated in the bacterial chromosome or in a plasmid, can be activated under stress, leading to the production of virions. These particles leave the host cell and can then harm and selectively kill other bacteria (79, 83).

The final category is that of chemical warfare (79, 80). Among chemical weapons, the most extensively studied are antimicrobial compounds or toxins (78), which usually target conserved pathways in the competitor, like translation and transcription machineries, the

cell envelope and the cytoskeleton (79). Central to that idea is the need for self and kin protection against these toxins, which is usually achieved by the production of immunity proteins (84, 85). Additionally, even though small toxins can be secreted and then diffuse across the target's membrane, larger toxins require specific mechanisms or machineries in order to be delivered into the target cell. Some examples of delivery machineries are contact-dependent growth inhibition (CDI) systems (86). These are two-partner secretion systems, where a very long protein exports the toxin tip out of the cell (87). The tip binds to a specific receptor on the target cell, is cleaved (88) and subsequently translocated into the cell (89). Other examples of toxin-delivery machineries are the so-called secretion systems. These systems and their secreted molecules and substrates can play essential roles in a variety of physiological bacterial processes but also in their interplay with the environment. So far there have been ten bacterial secretion systems recognized (90, 91), with the type X having been identified only very recently (92). Among them, types IV, VI and VII have been found to mediate interbacterial competition (79, 84, 93–97).

1.4.1 The type VI secretion system (T6SS)

Among these different bacterial molecular weapons, the T6SS is of particular interest to our research, since it is the only clear mechanism for bacterial antagonism found in *V. cholerae*, at least in well-studied pandemic strains (79). Furthermore, the three main gene clusters responsible for T6SS production in *V. cholerae* are found in all strains sequenced to date (pathogenic and environmental isolates) (98). This indicates that the machinery might be useful in both the environment as in a disease context.

Found in around one quarter of all Gram-negative bacteria (99), the T6SS resembles an inverted bacteriophage tail (100–102) that delivers effector molecules into target cells

(103, 104). The T6SS is composed of many proteins that anchor on a structure that spans both the inner and outer membranes of the cell (Figure 1.4). Proteins forming a baseplate along with the VgrG/PAAR spike are docked onto the membrane complex in a conformation that allows the assembly of an inner tube (formed by Hcp hexameric rings) and a contractile VipA/B sheath that encloses the Hcp tube. Sheath contraction propels the Hcp tube along with the VgrG/PAAR tip to the extracellular milieu or across a target cell's envelope (105, 106). The ClpV ATPase subsequently unfolds the contracted VipA/B sheath and recycles the subunits for a new assembly round of sheath extension and contraction (Figure 1.4)

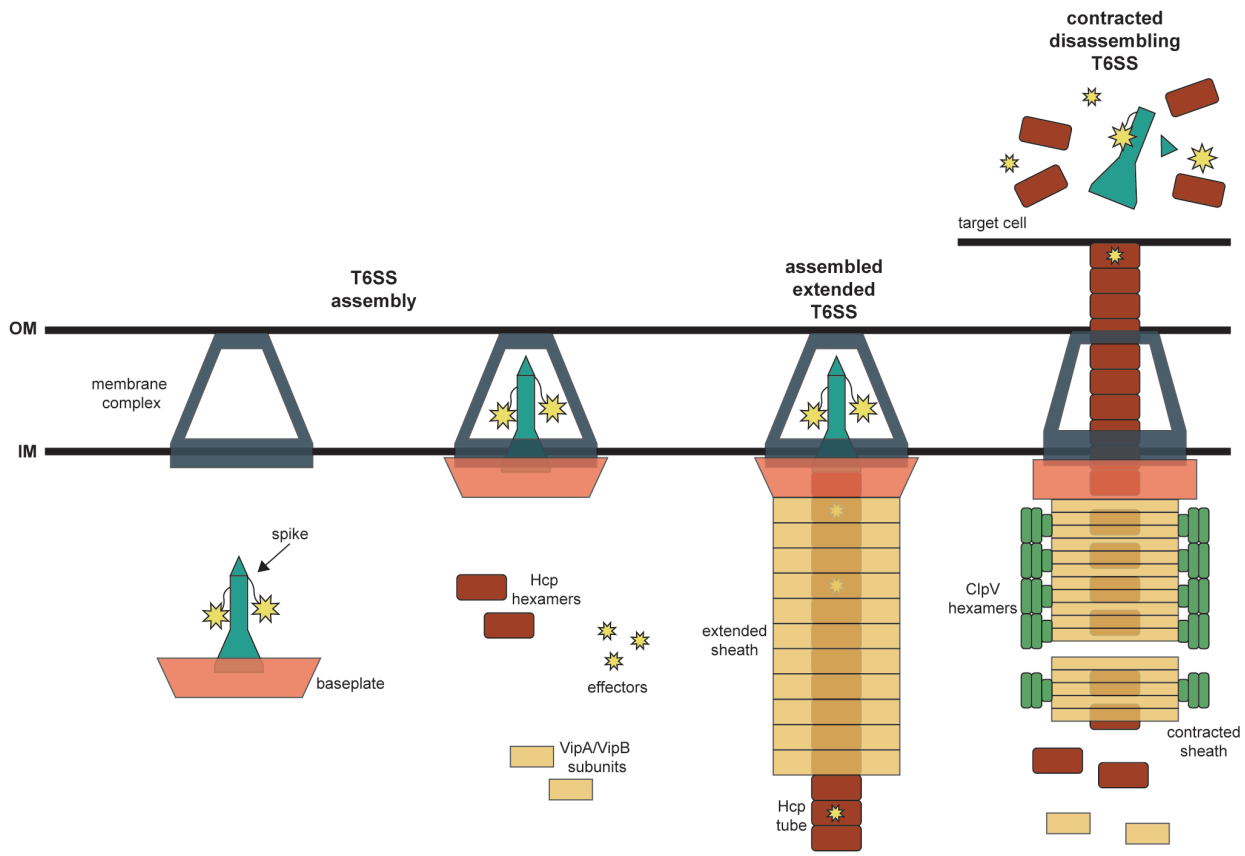


Figure 1.4 Functional dynamics of the T6SS.

T6SS assembly starts by the establishment of a membrane complex, to which a baseplate already assembled with a VgrG/PAAR spike are attached. This allows the assembly of a Hcp inner tube and of a VipA/B sheath around it. Toxic effectors can be loaded onto different structural components of the machinery or as C-terminal extensions of PAAR, VgrG or Hcp. Contraction of the sheath allows the Hcp tube and associated spike and effectors to be pushed out and potentially into a target cell, where the effectors will target conserved cellular components. The contracted VipA/B sheath is then disassembled by the ClpV ATPase, allowing recycling of the subunits for a new assembly line. Figure based on (105, 106).

Importantly, effector toxins are secreted by the T6SS either as C-terminal domains of VgrG, PAAR or Hcp (thereby called “evolved” versions of these proteins) or as cargos bound to structural proteins (such as Hcp or VgrG) (107). These effectors target conserved cellular components such as nucleic acids, membranes, peptidoglycan or the eukaryotic cytoskeleton (84, 108). Cells that produce the T6SS protect themselves from the toxic activity of these effectors by expressing cognate immunity proteins, which are usually encoded next to the effector genes, forming effector/ immunity (E/I) pairs (109–112). Besides self-protection, this system is employed as a mode of kin discrimination, as bacteria producing the same effector and immunity proteins are able to coexist (112).

1.4.1.1 Functionality of the T6SS

The T6SS was initially discovered in *V. cholerae* as a machinery used to fight predation by *D. discoideum* amoebae (69). In that same year, the T6SS of *Pseudomonas aeruginosa* was discovered and proposed to be involved in cystic fibrosis pathogenesis (113). The T6SS of other bacterial species were also demonstrated to be important in the interactions with eukaryotes. For example, the T6SS5 from the pseudomallei group of *Burkholderia* species (which harbor five different T6SS clusters) is essential for virulence in a mouse model but irrelevant for bacterial competition (96). The T6SS5 harbors VgrG5, which contains a C-terminal effector domain that is responsible for the formation of host multinucleated giant cells (114, 115). In the intracellular pathogen *Francisella*, a T6SS cluster found in the *Francisella* pathogenicity island (FPI) was shown to be pivotal for phagosomal escape and intracellular replication (116).

Despite these and other important examples of the T6SS function against eukaryotes, it has become increasingly clear that a primary purpose of the T6SS is to serve as an anti-

bacterial weapon (117). In fact, both species in which the T6SS was initially discovered were later shown to use the machinery to fight bacterial competitors (84, 97). Indeed, over the last decade a plethora of studies have illustrated the activity of the T6SS under controlled laboratory conditions, deepen our understanding of its structure, mode-of-action and regulation. Importantly, however, research has also been trying to investigate the T6SS under more natural conditions. For example, Speare and colleagues (118) observed that each crypt of the bobtail squid light organ is rarely colonized by more than one strain of *V. fischeri*. The authors beautifully demonstrated that this phenomenon was dependent on the T6SS, as certain strains use the machinery to outcompete others during early crypt colonization, leading to spatial separation of incompatible *V. fischeri* in different crypts.

The T6SS has been recently investigated in the context of the microbiota and especially the commensals of the mammalian gut. Bacteroidales are the most abundant and stable bacterial order in the human colon microbiota (119). More than half of the sequenced Bacteroidales strains in the human gut contain T6SS loci, which can be used for different competitive purposes (120–122). Furthermore, “colonization resistance” refers to the ability of a healthy microbiota to prevent the establishment of pathogens. The microbiota can achieve that by different mechanisms, such as stimulation of the host immune system, production of mucus, improvement of epithelial barrier integrity, competition for nutrients and metals, production of metabolites, direct inhibition, etc (123–126). Importantly, however, some pathogens have been shown to overcome such colonization resistance by using antagonistic strategies such as the T6SS. Recent work has indicated that pathogens such as *Salmonella enterica* serovar Typhimurium, *Shigella sonnei* and *V. cholerae* employ their T6SS to outcompete microbiota members *in vivo* (127–129) and for niche occupancy (127), even though these studies were mainly performed in animals that had their natural microbiota depleted by antibiotics.

1.4.1.2 T6SS clusters and regulation in pandemic *V. cholerae*

The T6SS of pandemic *V. cholerae* is encoded by four gene clusters: the large cluster and three auxiliary clusters (Aux 1-3) (Figure 1.5) (130). The large cluster contains mainly genes encoding structural components of the machinery. It also harbors a gene encoding one of the effectors, namely an evolved VgrG3 protein with a C-terminal domain with lysozyme activity, and the corresponding immunity protein TsiV3 (110, 131). The large cluster also contains an activator of RpoN (σ54) called VasH, which induces the expression of Aux 1 and 2 (132). These auxiliary clusters are similar in architecture, each one with a copy of the genes encoding Hcp and VgrG. Besides its structural role as part of the spike, VgrG1 encoded in Aux1 also contains an actin-crosslinking domain effector (70). In addition, both Aux1 and 2 carry a gene encoding an adaptor protein (Tap1 or VasW (133, 134)) and an E/I pair, namely the bifunctional lipase TseL and immunity TsiV1 in Aux1 (111, 131, 135) and the pore-forming toxin VasX and immunity TsiV2 in Aux2 (108, 136, 137). Importantly, the large and Aux clusters 1 and 2 are conserved among *V. cholerae* strains. Aux3, on the other hand, is largely enriched in pandemic *V. cholerae* strains, while a mobile and prophage-like form of this cluster can be sporadically found in environmental isolates (138). Aux3 contains a second copy of PAAR (the first copy is found in the large cluster) (139) and the E/I pair TseH/I (140, 141).

The T6SS of pandemic *V. cholerae* strains is tightly regulated and kept silent under laboratory conditions, though it can be activated by two independent signaling pathways. *V. cholerae* is commonly found associated to chitin surfaces in the aquatic environment (16). Besides serving as a food source, chitin also induces the physiological state of natural competence for transformation in this bacterium. This induction happens when cells reach high cell density (HCD) while growing on chitin and leads to the production of the competence activators TfoX and QstR (142–144). Importantly, the T6SS is concomitantly

produced under these conditions, therefore allowing the uptake of DNA released from killed non-kin neighboring bacteria (145). However, *V. cholerae* also contains another TfoX-like protein called TfoY (146), which has no role in natural competence for transformation (132). Consequently, TfoY works independently of TfoX, chitin or QS but instead is produced as a response to low intracellular c-di-GMP levels, which is measured post transcriptionally by a riboswitch that is located at the 5'-end of *tfoY* on the mRNA (132, 147).

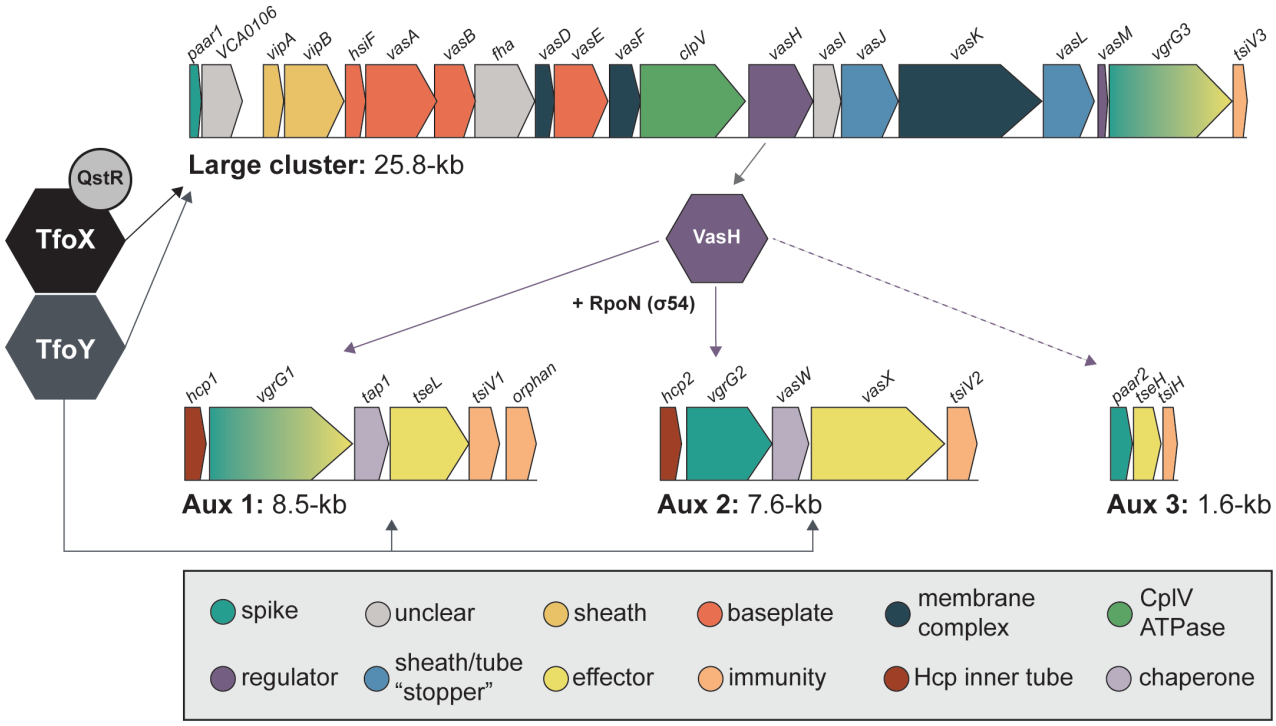


Figure 1.5 The T6SS clusters and regulation in 7th pandemic *V. cholerae*.

7th pandemic *V. cholerae* strains contain one T6SS, which is encoded by a large cluster and three auxiliary clusters. The large cluster contains genes encoding the main structural proteins of the machinery, as well as VasH, which acts as an activator of RpoN (σ54) and subsequently induces the expression of the auxiliary clusters 1 and 2. Aux 1 and 2 contain genes encoding Hcp and VgrG. All clusters, including Aux 3, contain an effector/immunity pair encoded by genes at their 3' end. The T6SS in pandemic strains can be activated by (i) TfoX/QstR when cells are grown to high cell density on chitin, or through (ii) TfoY, when c-di-GMP levels are low. TfoY also induces expression of the effector/immunity gene pairs in the Aux1-2 independently of VasH. Figure based on (132).

In sharp contrast to what is seen in pandemic strains, environmental isolates and non-pandemic toxigenic *V. cholerae* keep their T6SS constantly active, despite the

associated energy cost (51, 69, 101, 130, 148, 149). The mechanism behind constitutive T6SS activity in *V. cholerae* is currently not understood.

1.5 THESIS OUTLINE

Appropriate and refined public health measures for the control of infectious diseases are based on a well-rooted understanding of the ecology and evolution related to pathogen emergence and behavior. In the case of *V. cholerae*, strains causing the current 7th pandemics form a very successful lineage that is derived from a common ancestor in the Bay of Bengal in the 1950s, and which spread transcontinentally in three waves of transmission (28, 29). Importantly, single nucleotide polymorphism (SNP) analyses using 154 genomes of 7th pandemic *V. cholerae* strains demonstrated that the main source of variability among these strains came from continual local evolution in the Bay of Bengal (29). Furthermore, cholera is a highly seasonal disease, especially in endemic areas. Cholera peaks are generally observed before and after annual monsoons in Bangladesh (150, 151), and high *V. cholerae* concentrations are observed in environmental samples during periods with higher surface water temperatures and during zooplankton blooms (16, 151, 152). In order to comprehend the emergence and success of the pandemic lineage, it is therefore important that we understand *V. cholerae*'s behavior in the aquatic environment (1, 16, 18, 153). Moreover, as the coincidental evolution hypothesis states, many strategies and virulence factors employed by successful pathogens actually arose as environmentally-driven adaptations (3). Possible explanations for the success of the pandemic lineage could therefore come from studies comparing strategies employed by pandemic and environmental *V. cholerae* when dealing with environmental pressures. In that sense, this doctoral thesis aimed at investigating relevant aspects of *V. cholerae*'s biology in a

comparative framework, where 7th pandemic strains and a panel of fifteen environmental isolates collected in California in 2004 (154) were evaluated.

Chapter 2 of this thesis aimed at establishing a microscopy-based approach to visualize the interactions between *V. cholerae* and *D. discoideum* amoebae, initially focusing on the role of the T6SS. As previously mentioned, these amoebae were used in the initial study that led to the discovery of the T6SS (69), which was shown to deliver an actin-crosslinking effector into the grazers (70). While these experiments were all done on agar plates in so-called in plaque assays, the visualization of this killing mechanism was lacking, which we were hoping to develop using a microscopy-based approach. Interestingly, we uncovered an unexpected phenotype whereby *D. discoideum* are rapidly killed by *V. cholerae* in a so far unknown manner (independently of T6SS). Importantly, the induction of this phenotype varied between pandemic and environmental strains.

Considering how underappreciated the environmental *V. cholerae* diversity is, chapter 3 of this thesis was a “discovery expedition” into the genomes of those fifteen environmental isolates. We performed long-read PacBio-based whole genome sequencing and *de novo* assembly of these isolates and characterized the genomic islands that these strains carried. In sharp contrast to the considerably conserved mobilome from 7th pandemic strains, we uncovered a remarkably diverse horizontally transferable gene pool in these environmental isolates. We classified the environmental genomic islands according to their main genetic content into five categories: (i) antiphage defense systems; (ii) bacterial antagonism weapons; (iii) cell appendages; (iv) metabolism-related pathways; and (v) host-interaction systems. Transcriptional profiling (RNA sequencing) at both low and high cell density was employed to evaluate the expression level of the genes carried by these islands. This descriptive work will be supported by experimental data in the future.

Chapter 4 of this thesis is a published manuscript (155) in which we focused on understanding the differences in bacterial competition and predation evasion strategies used by environmental and pandemic *V. cholerae*. We observed that environmental *V. cholerae*, contrary to pandemic strains, constitutively secrete the pore-forming toxin hemolysin. However, this strategy was apparently not relevant as an anti-predation tool against *D. discoideum* grazing under the tested conditions. Furthermore, we observed that all environmental strains kept their T6SS constantly active, in contrast to the T6SS-silent status of pandemic strains. However, while all of these strains used their T6SS to kill *E. coli*, only a subset of them could use the machinery against protozoan grazers. We performed a full typing of the effector and immunity genes carried by these strains in all the T6SS clusters, which allowed us to pinpoint the specific anti-eukaryotic effector that is present in the specific subset of strains. Furthermore, we conducted pairwise killing experiments among all fifteen environmental isolates, which informed us on the degree of T6SS immunity polymorphism that still allows protection.

Chapter 5 aimed at investigating the phenomenon of constitutive T6SS activity in non-pandemic *V. cholerae* strains. While the tightly controlled T6SS activity from pandemic strains is quite well described (144, 145, 156), the highly costly constitutive T6SS production in non-pandemic isolates is not well understood. To understand the underlying mechanism, we constructed a library of 800 hybrid clones using a transformation-based cross of a pandemic and non-pandemic strain. These mosaic genome-carrying strains were then screened for their T6SS activity, which allowed us to find a causative SNP that controls T6SS activity in *V. cholerae*. We tackled many aspects of the SNP-derived T6SS regulation, which lead us to the proposition of a preliminary working model.

Chapter 6 of this thesis describes my contributions to other studies during my PhD. Specifically, I contributed the T6SS-related *E. coli* killing and microscopy data, and

constructed several mutants for the paper that described the interaction between *V. cholerae* and *A. castellanii* (74). Furthermore, I collaborated with a visiting PhD student from Carmen Amaro's lab from University of Valencia. This project aimed at investigating the interplay of *V. vulnificus* with *D. discoideum* using imaging approaches. Finally, I collaborated on a project that aimed to image fluorescently labelled *V. cholerae* in a minigut organoid system that was developed by the Lütolf lab (EPFL).

To conclude, this doctoral thesis provided several novel findings related to the diversity of environmental *V. cholerae* and their commonalities and differences when compared to pandemic *V. cholerae* patient isolates. The results from this work contribute to our understanding of *V. cholerae*'s biology, especially in the realm of the evolution of the pandemic lineage. Finally, the projects I initiated as part of this thesis created important new questions, as is common in science, and will open up new routes of investigation that can be pursued in the future.

1.6 REFERENCES

1. Martínez,J.L. (2013) Bacterial pathogens: from natural ecosystems to human hosts. *Environ. Microbiol.*, **15**, 325–333.

2. Martínez,J.L. (2014) Short-sighted evolution of bacterial opportunistic pathogens with an environmental origin. *Front. Microbiol.*, **5**, 1–4.

3. Levin,B.R. and Edén,C.S. (1990) Selection and evolution of virulence in bacteria: an ecumenical excursion and modest suggestion. *Parasitology*, **100**, S103–S115.

4. Adiba,S., Nizak,C., van Baalen,M., Denamur,E. and Depaulis,F. (2010) From grazing resistance to pathogenesis: The coincidental evolution of virulence factors. *PLoS One*, **5**, e11882.

5. Diard,M. and Hardt,W.D. (2017) Evolution of bacterial virulence. *FEMS Microbiol. Rev.*, **41**, 679–697.

6. Erken,M., Lutz,C. and McDougald,D. (2013) The rise of pathogens: predation as a factor driving the evolution of human pathogens in the environment. *Microb. Ecol.*, **65**, 860–868.

7. Steinberg,K.M. and Levin,B.R. (2007) Grazing protozoa and the evolution of the *Escherichia coli* O157:H7 Shiga toxin-encoding prophage. *Proc. R. Soc. B Biol. Sci.*, **274**, 1921–1929.

8. Brüssow,H. (2007) Bacteria between protists and phages: from antipredation strategies to the evolution of pathogenicity. *Mol. Microbiol.*, **65**, 583–589.

9. Dikow,R.B. and Smith,W.L. (2013) Genome-level homology and phylogeny of Vibrionaceae (Gammaproteobacteria: Vibrionales) with three new complete genome sequences. *BMC Microbiol.*, **13**.

10. Boyd,E.F., Carpenter,M.R., Chowdhury,N., Cohen,A.L., Haines-Menges,B.L., Kalburge,S.S., Kingston,J.J., Lubin,J.B., Ongagna-Yhombi,S.Y. and Whitaker,W.B. (2015) Post-genomic analysis of members of the family Vibrionaceae. *Microbiol. Spectr.*, **3**, VE-0009-2014.

11. Ali,M., Nelson,A.R., Lopez,A.L. and Sack,D.A. (2015) Updated global burden of cholera in endemic countries. *PLoS Negl. Trop. Dis.*, **9**, e0003832.

12. Faruque,S.M., Albert,M.J. and Mekalanos,J.J. (1998) Epidemiology, genetics, and ecology of toxigenic *Vibrio cholerae*. *Microbiol. Mol. Biol. Rev.*, **62**, 1301–14.

13. Kirn,T.J., Jude,B.A. and Taylor,R.K. (2005) A colonization factor links *Vibrio cholerae* environmental survival and human infection. *Nature*, **438**, 863–866.

14. Cottingham,K.L., Chiavelli,D.A. and Taylor,R.K. (2003) Environmental microbe and human pathogen: the ecology and microbiology of *Vibrio cholerae*. *Front Ecol Environm*, **1**, 80–86.

15. Tamplin,M.L., Gauzens,A.L., Huq,A., Sack,D.A. and Colwell,R.R. (1990) Attachment of *Vibrio cholerae* serogroup O1 to zooplankton and phytoplankton of Bangladesh waters. *Appl. Environ. Microbiol.*, **56**, 1977–1980.

16. Lipp,E.K., Huq,A. and Colwell,R.R. (2002) Effects of global climate on infectious disease: the cholera model. *Clin. Microbiol. Rev.*, **15**, 757–770.

17. Clemens,J.D., Nair,G.B., Ahmed,T., Qadri,F. and Holmgren,J. (2017) Cholera. *Lancet*, **390**, 1539–1549.

18. Colwell,R.R. (1996) Global climate and infectious disease: the cholera paradigm. *Science*, **274**, 2025–2031.

19. Blokesch,M. (2014) The lifestyle of *Vibrio cholerae* fosters gene transfers. *Microbe*, **9**, 64–70.

20. Almagro-Moreno,S., Pruss,K. and Taylor,R.K. (2015) Intestinal colonization dynamics of *Vibrio cholerae*. *PLoS Pathog.*, **11**, e1004787.

21. Butler,S.M. and Camilli,A. (2005) Going against the grain: chemotaxis and infection in *Vibrio cholerae*. *Nat. Rev. Microbiol.*, **3**, 611–620.

22. Cash,R.A., Music,S.I., Joseph,P., Snyder,M.J., Wenzel,R.P. and Hornick,R.B. (1974) Response of man to infection with *Vibrio cholerae*. I. Clinical, serologic, and bacteriologic response to a known inoculum. *J. Infect. Dis.*, **129**, 45–52.

23. Taylor,R.K., Miller,V.L., Furlong,D.B. and Mekalanos,J.J. (1987) Use of phoA gene fusions to identify a pilus colonization factor coordinately regulated with cholera toxin. *Proc. Natl. Acad. Sci.*, **84**, 2833–2837.

24. Kirn,T.J., Lafferty,M.J., Sandoe,C.M.P. and Taylor,R.K. (2000) Delineation of pilin domains required for bacterial association into microcolonies and intestinal colonization by *Vibrio cholerae*. *Mol. Microbiol.*, **35**, 896–910.

25. Waldor,M.K. and Mekalanos,J.J. (1996) Lysogenic conversion by a filamentous phage encoding cholera toxin. *Science*, **272**, 1910–1913.

26. Wernick,N.L.B., Chinnapen,D.J.F., Cho,J.A. and Lencer,W.I. (2010) Cholera toxin: An intracellular journey into the cytosol by way of the endoplasmic reticulum. *Toxins (Basel).*, **2**, 310–325.

27. Johnson,J.A., Salles,C.A., Panigrahi,P., Albert,M.J., Wright,A.C., Johnson,R.J. and Morris,J.G. (1994) *Vibrio cholerae* O139 synonym Bengal is closely related to *Vibrio cholerae* El Tor but has important differences. *Infect. Immun.*, **62**, 2108–2110.

28. Chun,J., Grim,C.J., Hasan,N.A., Lee,J.H., Choi,S.Y., Haley,B.J., Taviani,E., Jeon,Y.-S., Kim,D.W., Lee,J.-H., *et al.* (2009) Comparative genomics reveals mechanism for short-term and long-term clonal transitions in pandemic *Vibrio cholerae*. *Proc. Natl. Acad. Sci.*, **106**, 15442–15447.

29. Mutreja,A., Kim,D.W., Thomson,N.R., Connor,T.R., Lee,J.H., Kariuki,S., Croucher,N.J., Choi,S.Y., Harris,S.R., Lebens,M., *et al.* (2011) Evidence for several waves of global transmission in the seventh cholera pandemic. *Nature*, **477**, 462–465.

30. Weill,F.-X., Domman,D., Njamkepo,E., Tarr,C., Rauzier,J., Fawal,N., Keddy,K.H., Salje,H., Moore,S., Mukhopadhyay,A.K., *et al.* (2017) Genomic history of the seventh pandemic of cholera in Africa. *Science*, **358**, 785–789.

31. Domman,D., Quilici,M.-L., Dorman,M.J., Njamkepo,E., Mutreja,A., Mather,A.E., Delgado,G., Morales-Espinosa,R., Grimont,P.A.D., Lizarraga-Partida,M.L., *et al.* (2017) Integrated view of *Vibrio cholerae* in the Americas. *Science*, **358**, 789–793.

32. Doron,S., Melamed,S., Ofir,G., Leavitt,A., Lopatina,A., Keren,M., Amitai,G. and Sorek,R. (2018) Systematic discovery of antiphage defense systems in the microbial pangenome. *Science*, **359**, eaar4120.

33. Jermyn,W.S. and Boyd,E.F. (2002) Characterization of a novel *Vibrio* pathogenicity island (VPI-2) encoding neuraminidase (nanH) among toxigenic *Vibrio cholerae* isolates. *Microbiology*, **148**, 3681–3693.

34. Murphy,R.A. and Boyd,E.F. (2008) Three pathogenicity islands of *Vibrio cholerae* can excise from the chromosome and form circular intermediates. *J. Bacteriol.*, **190**, 636–647.

35. Dziejman,M., Balon,E., Boyd,D., Fraser,C.M., Heidelberg,J.F. and Mekalanos,J.J. (2002) Comparative genomic analysis of *Vibrio cholerae*: Genes that correlate with cholera endemic and pandemic disease. *Proc. Natl. Acad. Sci.*, **99**, 1556–1561.

36. O'Shea,Y.A., Finnan,S., Reen,F.J., Morrissey,J.P., O'Gara,F. and Boyd,E.F. (2004) The *Vibrio* seventh pandemic island-II is a 26.9 kb genomic island present in *Vibrio cholerae* El Tor and O139 serogroup isolates that shows homology to a 43.4 kb genomic island in *V. vulnificus*. *Microbiology*, **150**, 4053–4063.

37. Davies,B.W., Bogard,R.W., Young,T.S. and Mekalanos,J.J. (2012) Coordinated regulation of accessory genetic elements produces cyclic di-nucleotides for *V. cholerae* virulence. *Cell*, **149**, 358–370.

38. Cohen,D., Melamed,S., Millman,A., Shulman,G., Oppenheimer-Shaanan,Y., Kacen,A., Doron,S.,

Amitai,G. and Sorek,R. (2019) Cyclic GMP–AMP signalling protects bacteria against viral infection. *Nature*, **574**, 691–695.

39. Morehouse,B.R., Govande,A.A., Millman,A., Keszei,A.F.A., Lowey,B., Ofir,G., Shao,S., Sorek,R. and Kranzusch,P.J. (2020) STING cyclic dinucleotide sensing originated in bacteria. *Nature*, **586**, 429–433.

40. Le Roux,F. and Blokesch,M. (2018) Eco-evolutionary dynamics linked to horizontal gene transfer in Vibrios. *Annu. Rev. Microbiol.*, **72**, 89–110.

41. Dorman,M.J., Domman,D., Poklepovich,T., Tolley,C., Zolezzi,G., Kane,L., Viñas,M.R., Panagópulo,M., Moroni,M., Binsztein,N., *et al.* (2020) Genomics of the Argentinian cholera epidemic elucidate the contrasting dynamics of epidemic and endemic *Vibrio cholerae*. *Nat. Commun.*, **11**, 4918.

42. Robins,W.P. and Mekalanos,J.J. (2014) Genomic science in understanding cholera outbreaks and evolution of *Vibrio cholerae* as a human pathogen. In Nair,G., Takeda,Y. (eds), *Cholera Outbreaks. Current Topics in Microbiology and Immunology, vol 379.* Springer, Berlin, Heidelberg, pp. 211–229.

43. Aldova,E., Laznickova,K., Stepankova,E. and Lietava,J. (1968) Isolation of nonagglutinable vibrios from an enteritis outbreak in Czechoslovakia. *J. Infect. Dis.*, **118**, 25–31.

44. Boyd,E.F. and Waldor,M.K. (2002) Evolutionary and functional analyses of variants of the toxin-coregulated pilus protein TcpA from toxigenic *Vibrio cholerae* non-O1/nonO139 serogroup isolates. *Microbiology*, **148**, 1655–1666.

45. Haley,B.J., Choi,S.Y., Grim,C.J., Onifade,T.J., Cinar,H.N., Tall,B.D., Taviani,E., Hasan,N.A., Abdullah,A.H., Carter,L., *et al.* (2014) Genomic and phenotypic characterization of *Vibrio cholerae* non-O1 isolates from a US gulf coast cholera outbreak. *PLoS One*, **9**, e86264.

46. Faruque,S.M., Chowdhury,N., Kamruzzaman,M., Dziejman,M., Rahman,M.H., Sack,D.A., Nair,G.B. and Mekalanos,J.J. (2004) Genetic diversity and virulence potential of environmental *Vibrio cholerae* population in a cholera-endemic area. *Proc. Natl. Acad. Sci.*, **101**, 2123–2128.

47. Shapiro,B.J., Levade,I., Kovacikova,G., Taylor,R.K. and Almagro-Moreno,S. (2016) Origins of pandemic *Vibrio cholerae* from environmental gene pools. *Nat. Microbiol.*, **2**, 16240.

48. Rivera,I.N.G., Chun,J., Huq,A., Sack,R.B. and Colwell,R.R. (2001) Genotypes associated with virulence in environmental isolates of *Vibrio cholerae*. *Appl. Environ. Microbiol.*, **67**, 2421–2429.

49. Gennari,M., Ghidini,V., Caburlotto,G. and Lleo,M.M. (2012) Virulence genes and pathogenicity islands in environmental *Vibrio* strains nonpathogenic to humans. *FEMS Microbiol. Ecol.*, **82**, 563–573.

50. Faruque,S.M., Kamruzzaman,M., Meraj,I.M., Chowdhury,N., Nair,G.B., Sack,R.B., Colwell,R.R. and Sack,D.A. (2003) Pathogenic potential of environmental *Vibrio cholerae* strains carrying genetic variants of the toxin-coregulated pilus pathogenicity island. *Infect. Immun.*, **71**, 1020–1025.

51. Bernardy,E.E., Turnsek,M.A., Wilson,S.K., Tarr,C.L. and Hammer,B.K. (2016) Diversity of clinical and environmental isolates of *Vibrio cholerae* in natural tansformation and contact-dependent bacterial killing indicative of type VI secretion system activity. *Appl. Environ. Microbiol.*, **82**, 2833–2842.

52. Matz,C. and Kjelleberg,S. (2005) Off the hook - How bacteria survive protozoan grazing. *Trends Microbiol.*, **13**, 302–307.

53. Hahn,M.W. and Höfle,M.G. (2001) Grazing of protozoa and its effect on populations of aquatic bacteria. *FEMS Microbiol. Ecol.*, **35**, 113–121.

54. Pernthaler,J. (2005) Predation on prokaryotes in the water column and its ecological implications. *Nat. Rev. Microbiol.*, **3**, 537–546.

55. Justice,S.S., Hunstad,D.A., Cegelski,L. and Hultgren,S.J. (2008) Morphological plasticity as a bacterial survival strategy. *Nat. Rev. Microbiol.*, **6**, 162–168.

56. Matz,C., McDougald,D., Moreno,A.M., Yung,P.Y., Yildiz,F.H. and Kjelleberg,S. (2005) Biofilm formation and phenotypic variation enhance predation-driven persistence of *Vibrio cholerae*. *Proc. Natl. Acad.*

Sci., **102**, 16819–16824.

57. Sun,S., Noorian,P. and McDougald,D. (2018) Dual role of mechanisms involved in resistance to predation by protozoa and virulence to humans. *Front. Microbiol.*, **9**, 1017.

58. Haas,A. (2007) The phagosome: Compartment with a license to kill. *Traffic*, **8**, 311–330.

59. Scheid,P. (2014) Relevance of free-living amoebae as hosts for phylogenetically diverse microorganisms. *Parasitol. Res.*, **113**, 2407–2414.

60. Strassmann,J.E. and Shu,L. (2017) Ancient bacteria–amoeba relationships and pathogenic animal bacteria. *PLoS Biol.*, **15**, e2002460.

61. Hilbi,H., Weber,S.S., Ragaz,C., Nyfeler,Y. and Urwyler,S. (2007) Environmental predators as models for bacterial pathogenesis. *Environ. Microbiol.*, **9**, 563–575.

62. Casadevall,A. (2008) Evolution of intracellular pathogens. *Annu. Rev. Microbiol.*, **62**, 19–33.

63. Brown,M. and Barker,J. (1999) Unexplored reservoirs of pathogenic bacteria: protozoa and biofilms. *Trends Microbiol.*, **7**, 46–50.

64. Greub,G. and Raoult,D. (2004) Microorganisms resistant to free-living amoebae. *Clin. Microbiol. Rev.*, **17**, 413–433.

65. Molmeret,M., Horn,M., Wagner,M., Santic,M. and Kwaik,Y.A. (2005) Amoebae as training grounds for intracellular bacterial pathogens. *Appl. Environ. Microbiol.*, **71**, 20–28.

66. Escoll,P., Mondino,S., Rolando,M. and Buchrieser,C. (2015) Targeting of host organelles by pathogenic bacteria: A sophisticated subversion strategy. *Nat. Rev. Microbiol.*, **14**, 5–19.

67. Sun,S., Tay,Q.X.M., Kjelleberg,S., Rice,S.A. and McDougald,D. (2015) Quorum sensing-regulated chitin metabolism provides grazing resistance to *Vibrio cholerae* biofilms. *ISME J.*, **9**, 1812–1820.

68. Vaitkevicius,K., Lindmark,B., Ou,G., Song,T., Toma,C., Iwanaga,M., Zhu,J., Andersson,A., Hammarstrom,M.-L., Tuck,S., *et al.* (2006) A *Vibrio cholerae* protease needed for killing of *Caenorhabditis elegans* has a role in protection from natural predator grazing. *Proc. Natl. Acad. Sci.*, **103**, 9280–9285.

69. Pukatzki,S., Ma,A.T., Sturtevant,D., Krastins,B., Sarracino,D., Nelson,W.C., Heidelberg,J.F. and Mekalanos,J.J. (2006) Identification of a conserved bacterial protein secretion system in *Vibrio cholerae* using the *Dictyostelium* host model system. *Proc. Natl. Acad. Sci. U. S. A.*, **103**, 1528–33.

70. Pukatzki,S., Ma,A.T., Revel,A.T., Sturtevant,D. and Mekalanos,J.J. (2007) Type VI secretion system translocates a phage tail spike-like protein into target cells where it cross-links actin. *Proc. Natl. Acad. Sci.*, **104**, 15508–15513.

71. Ma,A.T., McAuley,S., Pukatzki,S. and Mekalanos,J.J. (2009) Translocation of a *Vibrio cholerae* type VI secretion effector requires bacterial endocytosis by host cells. *Cell Host Microbe*, **5**, 234–243.

72. Vezzulli,L., Pruzzo,C., Huq,A. and Colwell,R.R. (2010) Environmental reservoirs of *Vibrio cholerae* and their role in cholera. *Environ. Microbiol. Rep.*, **2**, 27–33.

73. Van Der Henst,C., Scrignari,T., Maclachlan,C. and Blokesch,M. (2016) An intracellular replication niche for *Vibrio cholerae* in the amoeba *Acanthamoeba castellanii*. *ISME J.*, **10**, 897–910.

74. Van der Henst,C., Vanhove,A.S., Drebes Dörr,N.C., Stutzmann,S., Stoudmann,C., Clerc,S., Scrignari,T., Maclachlan,C., Knott,G. and Blokesch,M. (2018) Molecular insights into *Vibrio cholerae*'s intra-amoebal host-pathogen interactions. *Nat. Commun.*, **9**, 3460.

75. Espinoza-Vergara,G., Noorian,P., Silva-Valenzuela,C.A., Raymond,B.B.A., Allen,C., Hoque,M.M., Sun,S., Johnson,M.S., Pernice,M., Kjelleberg,S., *et al.* (2019) *Vibrio cholerae* residing in food vacuoles expelled by protozoa are more infectious *in vivo*. *Nat. Microbiol.*, **4**, 2466–2474.

76. Flemming,H.C. and Wuertz,S. (2019) Bacteria and archaea on Earth and their abundance in biofilms. *Nat. Rev. Microbiol.*, **17**, 247–260.

77. Stubbendieck,R.M. and Straight,P.D. (2016) Multifaceted interfaces of bacterial competition. *J. Bacteriol.*, **198**, 2145–2155.

78. Hibbing,M.E., Fuqua,C., Parsek,M.R. and Peterson,S.B. (2010) Bacterial competition: surviving and thriving in the microbial jungle. *Nat. Rev. Microbiol.*, **8**, 15–25.

79. Granato,E.T., Meiller-Legrand,T.A. and Foster,K.R. (2019) The evolution and ecology of bacterial warfare. *Curr. Biol.*, **29**, R521–R537.

80. Peterson,S.B., Bertolli,S.K. and Mougous,J.D. (2020) The central role of interbacterial antagonism in bacterial fife. *Curr. Biol.*, **30**, R1203–R1214.

81. Scholl,D. (2017) Phage tail – like bacteriocins. *Annu. Rev. Virol.*, **4**, 453–467.

82. Ghequire,M.G.K. and De Mot,R. (2015) The tailocin tale: peeling off phage tails. *Trends Microbiol.*, **23**, 587–590.

83. Howard-Varona,C., Hargreaves,K.R., Abedon,S.T. and Sullivan,M.B. (2017) Lysogeny in nature: mechanisms, impact and ecology of temperate phages. *ISME J.*, **11**, 1511–1520.

84. Hood,R.D., Singh,P., Hsu,F.S., Güvener,T., Carl,M.A., Trinidad,R.R.S., Silverman,J.M., Ohlson,B.B., Hicks,K.G., Plemel,R.L., *et al.* (2010) A type VI secretion system of *Pseudomonas aeruginosa* targets a toxin to bacteria. *Cell Host Microbe*, **7**, 25–37.

85. Cascales,E., Buchanan,S.K., Duché,D., Kleanthous,C., Lloubès,R., Postle,K., Riley,M., Slatin,S. and Cavard,D. (2007) Colicin biology. *Microbiol. Mol. Biol. Rev.*, **71**, 158–229.

86. Aoki,S.K., Diner,E.J., De Roodenbeke,C.T.K., Burgess,B.R., Poole,S.J., Braaten,B.A., Jones,A.M., Webb,J.S., Hayes,C.S., Cotter,P.A., *et al.* (2010) A widespread family of polymorphic contact-dependent toxin delivery systems in bacteria. *Nature*, **468**, 439–442.

87. Ruhe,Z.C., Low,D.A. and Hayes,C.S. (2013) Bacterial contact-dependent growth inhibition. *Trends Microbiol.*, **21**, 230–237.

88. Ruhe,Z.C., Subramanian,P., Song,K., Nguyen,J.Y., Stevens,T.A., Low,D.A., Jensen,G.J. and Hayes,C.S. (2018) Programmed secretion arrest and receptor-triggered toxin export during antibacterial contact-dependent growth inhibition. *Cell*, **175**, 921–933.

89. Hayes,C.S., Koskiniemi,S., Ruhe,Z.C., Poole,S.J. and Low,D.A. (2014) Mechanisms and biological roles of contact- dependent growth inhibition systems. *Cold Spring Harb. Perspect. Med.*, **14**, a010025.

90. Costa,T.R.D., Felisberto-Rodrigues,C., Meir,A., Prevost,M.S., Redzej,A., Trokter,M. and Waksman,G. (2015) Secretion systems in Gram-negative bacteria: structural and mechanistic insights. *Nat. Rev. Microbiol.*, **13**, 343–359.

91. Green,E.R. and Mecsas,J. (2016) Bacterial secretion systems: an overview. *Microbiol. Spectr.*, **4**, VMBF-0012-2015.

92. Palmer,T., Finney,A.J., Saha,C.K., Atkinson,G.C. and Sargent,F. (2020) A holin/peptidoglycan hydrolase-dependent protein secretion system. *Mol. Microbiol.*, 10.1111/mmi.14599.

93. Souza,D.P., Oka,G.U., Alvarez-Martinez,C.E., Bisson-Filho,A.W., Dunger,G., Hobeika,L., Cavalcante,N.S., Alegria,M.C., Barbosa,L.R.S., Salinas,R.K., *et al.* (2015) Bacterial killing via a type IV secretion system. *Nat. Commun.*, **6**, 6453.

94. Cao,Z., Casabona,M.G., Kneuper,H., Chalmers,J.D. and Palmer,T. (2016) The type VII secretion system of Staphylococcus aureus secretes a nuclease toxin that targets competitor bacteria. *Nat. Microbiol.*, **2**.

95. Whitney,J.C., Peterson,S.B., Kim,J., Pazos,M., Verster,A.J., Radey,M.C., Kulasekara,H.D., Ching,M.Q., Bullen,N.P., Bryant,D., *et al.* (2017) A broadly distributed toxin family mediates contact-dependent

antagonism between gram-positive bacteria. *Elife*, **6**, e26938.

96. Schwarz,S., West,T.E., Boyer,F., Chiang,W.C., Carl,M.A., Hood,R.D., Rohmer,L., Tolker-Nielsen,T., Skerrett,S.J. and Mougous,J.D. (2010) Burkholderia type vi secretion systems have distinct roles in eukaryotic and bacterial cell interactions. *PLoS Pathog.*, **6**, e1001068.

97. MacIntyre,D.L., Miyata,S.T., Kitaoka,M. and Pukatzki,S. (2010) The *Vibrio cholerae* type VI secretion system displays antimicrobial properties. *Proc. Natl. Acad. Sci.*, **107**, 19520–19524.

98. Boyer,F., Fichant,G., Berthod,J., Vandenbrouck,Y. and 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**.

99. Ho,B.T., Dong,T.G. and Mekalanos,J.J. (2014) A view to a kill: the bacterial type VI secretion system. *Cell Host Microbe*, **15**, 9–21.

100. Leiman,P.G., Basler,M., Ramagopal,U.A., Bonanno,J.B., Sauder,J.M., Pukatzki,S., Burley,S.K., Almo,S.C. and Mekalanos,J.J. (2009) Type VI secretion apparatus and phage tail-associated protein complexes share a common evolutionary origin. *Proc. Natl. Acad. Sci. U. S. A.*, **106**, 4154–4159.

101. Basler,M., Pilhofer,M., Henderson,G.P., Jensen,G.J. and Mekalanos,J.J. (2012) Type VI secretion requires a dynamic contractile phage tail-like structure. *Nature*, **483**, 182–186.

102. Taylor,N.M.I., van Raaij,M.J. and Leiman,P.G. (2018) Contractile injection systems of bacteriophages and related systems. *Mol. Microbiol.*, **108**, 6–15.

103. Galán,J.E. and Waksman,G. (2018) Protein-injection machines in bacteria. *Cell*, **172**, 1306–1318.

104. Cianfanelli,F.R., Monlezun,L. and Coulthurst,S.J. (2016) Aim, load, fire: the type VI secretion system, a bacterial nanoweapon. *Trends Microbiol.*, **24**, 51–62.

105. Basler,M. (2015) Type VI secretion system: secretion by a contractile nanomachine. *Philos. Trans. R. Soc. B Biol. Sci.*, **370**, 20150021.

106. Cherrak,Y., Flaugnatti,N., Durand,E., Journet,L. and Cascales,E. (2019) Structure and regulation of the type VI secretion system. *Microbiol. Spectr.*, **7**, PSIB-0031-2019.

107. Hachani,A., Wood,T.E. and Filloux,A. (2016) Type VI secretion and anti-host effectors. *Curr. Opin. Microbiol.*, **29**, 81–93.

108. Russell,A.B., Peterson,S.B. and Mougous,J.D. (2014) Type VI secretion system effectors: poisons with a purpose. *Nat. Rev. Microbiol.*, **12**, 137–148.

109. Russell,A.B., Hood,R.D., Bui,N.K., Leroux,M., Vollmer,W. and Mougous,J.D. (2011) Type VI secretion delivers bacteriolytic effectors to target cells. *Nature*, **475**, 343–349.

110. Brooks,T.M., Unterweger,D., Bachmann,V., Kostiuk,B. and Pukatzki,S. (2013) Lytic activity of the *Vibrio cholerae* type VI secretion toxin VgrG-3 is inhibited by the antitoxin TsaB. *J. Biol. Chem.*, **288**, 7618–7625.

111. Dong,T.G., Ho,B.T., Yoder-Himes,D.R. and Mekalanos,J.J. (2013) Identification of T6SS-dependent effector and immunity proteins by Tn-seq in *Vibrio cholerae*. *Proc. Natl. Acad. Sci.*, **110**, 2623–2628.

112. Unterweger,D., Miyata,S.T., Bachmann,V., Brooks,T.M., Mullins,T., Kostiuk,B., Provenzano,D. and Pukatzki,S. (2014) The *Vibrio cholerae* type VI secretion system employs diverse effector modules for intraspecific competition. *Nat. Commun.*, **5**, 3549.

113. Mougous,J.D., Cuff,M.E., Raunser,S., Shen,A., Zhou,M., Gifford,C.A., Goodman,A.L., Joachimiak,G., Ordoñez,C.L., Lory,S., *et al.* (2006) A virulence locus of *Pseudomonas aeruginosa* encodes a protein secretion apparatus. *Science*, **312**, 1526–1530.

114. Toesca,I.J., French,C.T. and Miller,J.F. (2014) The type VI secretion system spike protein VgrG5 mediates membrane fusion during intercellular spread by pseudomallei group *Burkholderia* species.

Infect. Immun., **82**, 1436–1444.

115. Schwarz,S., Singh,P., Robertson,J.D., LeRoux,M., Skerrett,S.J., Goodlett,D.R., Eoin West,T. and Mougous,J.D. (2014) VgrG-5 is a *Burkholderia* type VI secretion system-exported protein required for multinucleated giant cell formation and virulence. *Infect. Immun.*, **82**, 1445–1452.

116. Clemens,D.L., Lee,B.Y. and Horwitz,M.A. (2018) The *Francisella* type VI secretion system. *Front. Cell. Infect. Microbiol.*, **8**, 121.

117. Hood,R.D., Peterson,S.B. and Mougous,J.D. (2017) From striking out to striking gold: discovering that type VI secretion targets bacteria. *Cell Host Microbe*, **21**, 286–289.

118. Speare,L., Cecere,A.G., Guckes,K.R., Smith,S., Wollenberg,M.S., Mandel,M.J., Miyashiro,T. and Septer,A.N. (2018) Bacterial symbionts use a type VI secretion system to eliminate competitors in their natural host. *Proc. Natl. Acad. Sci. U. S. A.*, **115**, E8528–E8537.

119. Faith,J.J., Guruge,J.L., Charbonneau,M., Subramanian,S., Seedorf,H., Goodman,A.L., Clemente,J.C., Knight,R., Heath,A.C., Leibel,R.L., *et al.* (2013) The long-term stability of the human gut microbiota. *Science*, **341**.

120. Coyne,M.J., Roelofs,K.G. and Comstock,L.E. (2016) Type VI secretion systems of human gut Bacteroidales segregate into three genetic architectures, two of which are contained on mobile genetic elements. *BMC Genomics*, **17**, 58.

121. Chatzidaki-Livanis,M., Geva-Zatorsky,N., Comstock,L.E. and Hooper,L. V. (2016) *Bacteroides fragilis* type VI secretion systems use novel effector and immunity proteins to antagonize human gut Bacteroidales species. *Proc. Natl. Acad. Sci. U. S. A.*, **113**, 3627–3632.

122. Verster,A.J., Ross,B.D., Radey,M.C., Bao,Y., Goodman,A.L., Mougous,J.D. and Borenstein,E. (2017) The landscape of type VI secretion across human gut microbiomes reveals its role in community composition. *Cell Host Microbe*, **22**, 411–419.

123. Leshem,A., Liwinski,T. and Elinav,E. (2020) Immune-microbiota interplay and colonization resistance in infection. *Mol. Cell*, **78**, 597–613.

124. Ducarmon,Q.R., Zwittink,R.D., Hornung,B.V.H., van Schaik,W., Young,V.B. and Kuijper,E.J. (2019) Gut microbiota and colonization resistance against bacterial enteric infection. *Microbiol. Mol. Biol. Rev.*, **83**, e00007-19.

125. Buffie,C.G. and Pamer,E.G. (2013) Microbiota-mediated colonization resistance against intestinal pathogens. *Nat. Rev. Immunol.*, **13**, 790–801.

126. Mullineaux-Sanders,C., Suez,J., Elinav,E. and Frankel,G. (2018) Sieving through gut models of colonization resistance. *Nat. Microbiol.*, **3**, 132–140.

127. Anderson,M.C., Vonaesch,P., Saffarian,A., Marteyn,B.S. and Sansonetti,P.J. (2017) *Shigella sonnei* encodes a functional T6SS used for interbacterial competition and niche occupancy. *Cell Host Microbe*, **21**, 769–776.

128. Zhao,W., Caro,F., Robins,W. and Mekalanos,J.J. (2018) Antagonism toward the intestinal microbiota and its effect on *Vibrio cholerae* virulence. *Science*, **359**, 210–213.

129. Sana,T.G., Flaugnatti,N., Lugo,K.A., Lam,L.H., Jacobson,A., Baylot,V., Durand,E., Journet,L., Cascales,E. and Monack,D.M. (2016) *Salmonella* Typhimurium utilizes a T6SS-mediated antibacterial weapon to establish in the host gut. *Proc. Natl. Acad. Sci.*, **113**, E5044–E5051.

130. Crisan,C. V and Hammer,B.K. (2020) The *Vibrio cholerae* type VI secretion system: toxins, regulators and consequences. *Environ. Microbiol.*, **22**, 4112–4122.

131. Zheng,J., Ho,B. and Mekalanos,J.J. (2011) Genetic analysis of anti-amoebae and anti-bacterial activities of the type VI secretion system in *Vibrio cholerae*. *PLoS One*, **6**, e23876.

132. Metzger,L.C., Stutzmann,S., Scrignari,T., Van der Henst,C., Matthey,N. and Blokesch,M. (2016)

Independent regulation of type VI secretion in *Vibrio cholerae* by TfoX and TfoY. *Cell Rep.*, **15**, 951–958.

133. Liang,X., Moore,R., Wilton,M., Wong,M.J.Q., Lam,L. and Dong,T.G. (2015) Identification of divergent type VI secretion effectors using a conserved chaperone domain. *Proc. Natl. Acad. Sci.*, **112**, 9106–9111.

134. Unterweger,D., Kostiuk,B., Otjengerdes,R., Wilton,A., Diaz-Satizabal,L. and Pukatzki,S. (2015) Chimeric adaptor proteins translocate diverse type VI secretion system effectors in *Vibrio cholerae*. *EMBO J.*, **34**, 2198–2210.

135. Russell,A.B., Leroux,M., Hathazi,K., Agnello,D.M., Ishikawa,T., Wiggins,P.A., Wai,S.N. and Mougous,J.D. (2013) Diverse type VI secretion phospholipases are functionally plastic antibacterial effectors. *Nature*, **496**, 508–512.

136. Miyata,S.T., Kitaoka,M., Brooks,T.M., McAuley,S.B. and Pukatzki,S. (2011) *Vibrio cholerae* requires the type VI secretion system virulence factor VasX to kill *Dictyostelium discoideum*. *Infect. Immun.*, **79**, 2941–2949.

137. Miyata,S.T., Unterweger,D., Rudko,S.P. and Pukatzki,S. (2013) Dual expression profile of type VI secretion system immunity genes protects pandemic *Vibrio cholerae*. *PLoS Pathog.*, **9**, e1003752.

138. Santoriello,F.J., Michel,L., Unterweger,D. and Pukatzki,S. (2020) Pandemic *Vibrio cholerae* shuts down site-specific recombination to retain an interbacterial defence mechanism. *Nat. Commun.*, **11**, 6246.

139. Shneider,M.M., Buth,S.A., Ho,B.T., Basler,M., Mekalanos,J.J. and Leiman,P.G. (2013) PAAR-repeat proteins sharpen and diversify the type VI secretion system spike. *Nature*, **500**, 350–353.

140. Hersch,S.J., Watanabe,N., Stietz,M.S., Manera,K., Kamal,F., Burkinshaw,B., Lam,L., Pun,A., Li,M., Savchenko,A., *et al.* (2020) Envelope stress responses defend against type six secretion system attacks independently of immunity proteins. *Nat. Microbiol.*, **5**, 706–714.

141. Altindis,E., Dong,T., Catalano,C. and Mekalanos,J.J. (2015) Secretome analysis of *Vibrio cholerae* type VI secretion system reveals a new effector-immunity pair. *MBio*, **6**, e00075-15.

142. Meibom,K.L., Blokesch,M., Dolganov,N.A., Wu,C.Y. and Schoolnik,G.K. (2005) Chitin induces natural competence in *Vibrio cholerae*. *Science*, **310**, 1824–1827.

143. Lo Scrudato,M. and Blokesch,M. (2013) A transcriptional regulator linking quorum sensing and chitin induction to render *Vibrio cholerae* naturally transformable. *Nucleic Acids Res.*, **41**, 3644–3658.

144. Jaskólska,M., Stutzmann,S., Stoudmann,C. and Blokesch,M. (2018) QstR-dependent regulation of natural competence and type VI secretion in *Vibrio cholerae*. *Nucleic Acids Res.*, **46**, 10619–10634.

145. Borgeaud,S., Metzger,L.C., Scrignari,T. and Blokesch,M. (2015) The type VI secretion system of *Vibrio cholerae* fosters horizontal gene transfer . *Science*, **347**, 63–67.

146. Pollack-Berti,A., Wollenberg,M.S. and Ruby,E.G. (2010) Natural transformation of *Vibrio fischeri* requires *tfoX* and *tfoY*. *Environ. Microbiol.*, **12**, 2302–2311.

147. Sudarsan,N., Lee,E.R., Weinberg,Z., Moy,R.H., Kim,J.N., Link,K.H. and Breaker,R.R. (2008) Riboswitches in eubacteria sense the second messenger cyclic Di-GMP. *Science*, **321**, 411–413.

148. Van der Henst,C., Vanhove,A.S., Drebes Dörr,N.C., Stutzmann,S., Stoudmann,C., Clerc,S., Scrignari,T., Maclachlan,C., Knott,G. and Blokesch,M. (2018) Molecular insights into *Vibrio cholerae*'s intra-amoebal host-pathogen interactions. *Nat. Commun.*, **9**, 3460.

149. Unterweger,D., Kitaoka,M., Miyata,S.T., Bachmann,V., Brooks,T.M., Moloney,J., Sosa,O., Silva,D., Duran-Gonzalez,J., Provenzano,D., *et al.* (2012) Constitutive type VI secretion system expression gives *Vibrio cholerae* intra- and interspecific competitive advantages. *PLoS One*, **7**, e48320.

150. Longini,I.M., Yunus,M., Zaman,K., Siddique,A.K., Sack,R.B. and Nizam,A. (2002) Epidemic and endemic cholera trends over a 33-year period in Bangladesh. *J. Infect. Dis.*, **186**, 246–251.

151. Emch,M., Feldacker,C., Sirajul,S.M. and Ali,M. (2008) Seasonality of cholera from 1974 to 2005: A review of global patterns. *Int. J. Health Geogr.*, **7**.

152. Lobitz,B., Beck,L., Huq,A., Wood,B., Fuchs,G., Faruque,A.S.G. and Colwell,R. (2000) Climate and infectious disease: use of remote sensing for detection of *Vibrio cholerae* by indirect measurement. *Proc. Natl. Acad. Sci. U. S. A.*, **97**, 1438–1443.

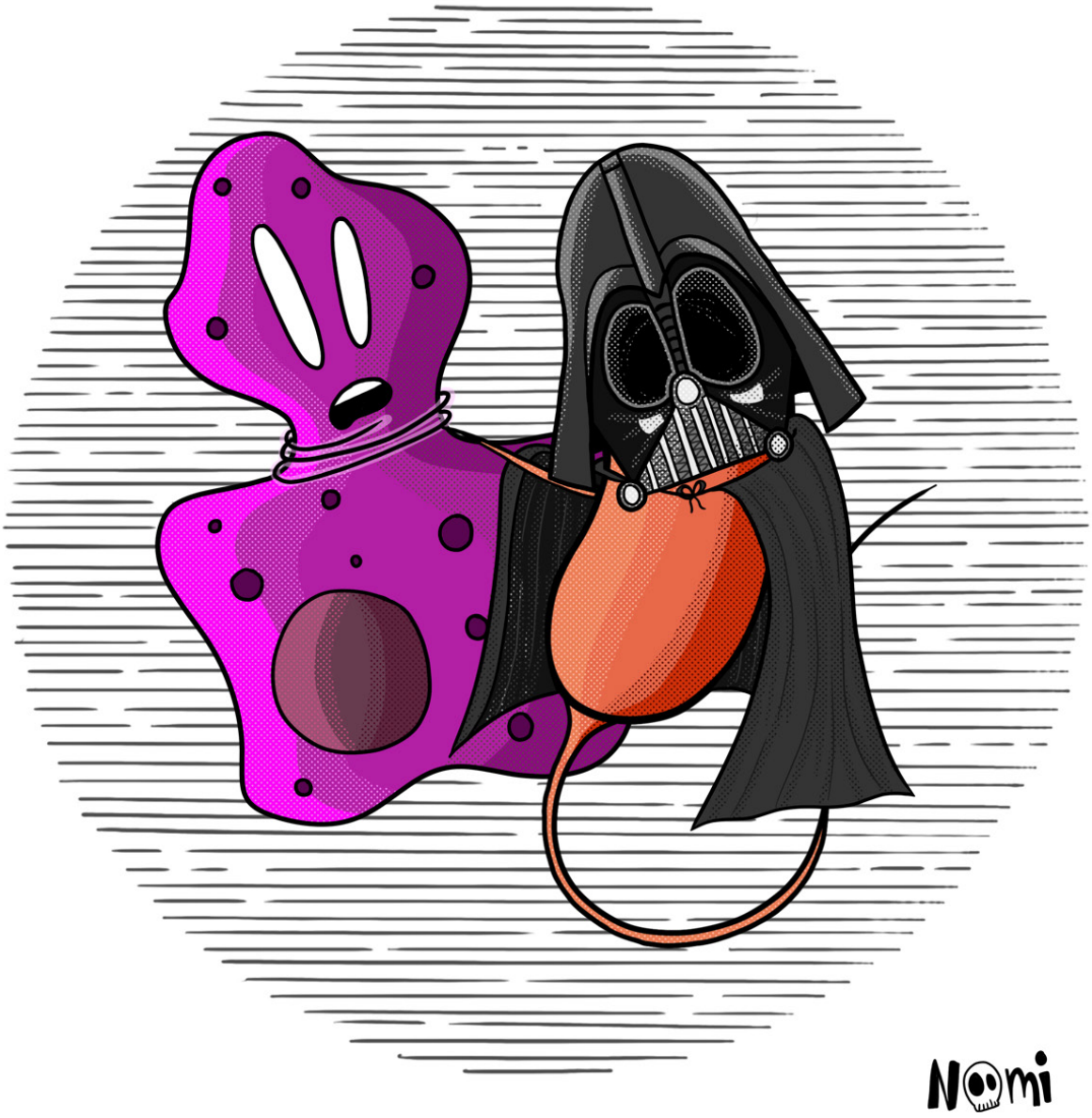
153. Alam,M., Kasan,N.A., Sadique,A., Bhuiyan,N.A., Ahmed,K.U., Nusrin,S., Nair,G.B., Siddique,A.K., Sack,R.B., Sack,D.A., *et al.* (2006) Seasonal cholera caused by *Vibrio cholerae* serogroups O1 and O139 in the coastal aquatic environment of Bangladesh. *Appl. Environ. Microbiol.*, **72**, 4096–4104.

154. Keymer,D.P., Miller,M.C., Schoolnik,G.K. and Boehm,A.B. (2007) Genomic and phenotypic diversity of coastal *Vibrio cholerae* strains is linked to environmental factors. *Appl. Environ. Microbiol.*, **73**, 3705–3714.

155. Drebes Dörr,N.C. and Blokesch,M. (2020) Interbacterial competition and anti-predatory behavior of environmental *Vibrio cholerae* strains. *Environ. Microbiol.*, **22**, 4485–4504.

156. Metzger,L.C. and Blokesch,M. (2016) Regulation of competence-mediated horizontal gene transfer in the natural habitat of *Vibrio cholerae*. *Curr. Opin. Microbiol.*, **30**, 1–7.

2 DEFENSE STRATEGIES OF *VIBRIO CHOLERAE* AGAINST AMOEBAL PREDATION



Drawing by Noémie Matthey, PhD

2.1 INTRODUCTION

As we seek ways to battle bacteria that cause disease, we tend to forget that these microorganisms have been involved in warfares of their own long before the first animals arose (1). Aside from abiotic challenges such as food availability, changes in temperature, pH and so on, bacteria encounter pressure from many biotic sources in the environment they inhabit. As microbes are commonly embedded in communities (2), competition for food and space is central for their survival while living with or outcompeting other bacteria. In this sense, a broad spectrum of antagonizing strategies has emerged during evolution, which were discussed in great detail in recent reviews such as (3, 4). Besides that, bacteria have faced predation burden from bacteriophages and eukaryotic grazers, most likely since billions of years (1, 5). With the onset of comparative genomics, there has been an explosion on *in silico* discoveries of new antiphage defense systems, many of which have been experimentally validated (though in surrogate hosts such as *Escherichia coli* or *Bacillus subtills*) (6, 7). On the other hand, defense mechanisms against protozoan grazers are often less predictable by genomics. As such grazers force bacteria into their food vacuole(s) by phagocytosis, bacterial defense mechanisms can occur extracellularly (before ingestion) or intracellularly (inside the vacuole). Extracellular adaptations include, among others, morphological plasticity to avoid or diminish grazing rate, escape from predation using motility strategies, secretion of exopolymers leading to biofilm formation and changes in surface components to avoid recognition by phagocytic receptors (1, 5, 8, 9).

Conversely, intracellular adaptations aim at dealing with the harsh conditions encountered in the digestive vacuoles. Such conditions include acidification, oxidative stress, nutrient deprivation, and the production of antimicrobial peptides and digestive enzymes. Bacteria have developed strategies to either resist these conditions or to kill predators after their ingestion through the release of toxins (10). In fact, some microbes

have adapted to hijack the phagosome-lysosome pathway and utilize it as a replication niche. In the environment, phagocytic amoebae can therefore serve as training grounds for pathogenicity (11, 12). Indeed, environmentally-driven adaptations have evolved to become mechanistically intricate and complex, giving rise to important facultative or obligate pathogens such as *Mycobacterium*, *Legionella*, *Chlamydia* and *Listeria* (1, 5, 12–18).

Clarifying the mechanisms that allow a shift from an environmental to a pathogenic lifestyle is critical for our better understanding of microbes that impact human health. *Vibrio cholerae*, a Gram-negative bacterium that is known for causing the deadly diarrheal disease cholera, is an excellent model to study this shift. Despite being notoriously known for the disease it causes, *V. cholerae* is actually a diverse species and is composed mainly of environmental strains that do not cause typical cholera-like gastroenteritis. Notably, only two (O1 and O139) out of more than 200 known serogroups have been responsible for the seven reported cholera pandemics that humankind has experienced since 1817 (19–22). These almost clonal pandemic strains have clear genomic features when compared to environmental isolates. Among those are the main cholera virulence factors-encoding genes (cholera toxin and toxin-coregulated pilus) on genomic islands/prophages (23, 24). Nonetheless, the acquisition of these islands is required but not sufficient for a *V. cholerae* strain to gain pandemic potential. Indeed, it is still not entirely clear which combination of factors drove the evolution of the most successful pandemic lineage of *V. cholerae*.

Besides having defined O-antigens (which specify the serogroups) and a characteristic set of genomic islands (25–28), pandemic *V. cholerae* strains also display distinct phenotypic features when compared to environmental isolates. For instance, they keep a tight regulation of important molecular weapons, such as the type VI secretion system (T6SS) and the pore-forming toxin hemolysin, while environmental isolates keep these machineries constitutively active (29–32). These differences could not be predicted

by genomics, as all *V. cholerae* strains are known to possess the genes/gene clusters that encode these features (33, 34). Importantly, the regulatory cascades that control the production of the T6SS and the hemolysin might have contributed to the emergence of the pandemic lineage. Contrary to this "on demand" production, environmental isolates likely benefit from maintaining their molecular weapons in a constantly active state (31).

As environmental adaptations might be coopted in a disease situation, we were interested in investigating which strategies *V. cholerae* could use as a defense against protozoan predation. Besides providing important information about its environmental lifestyle, this approach could potentially decipher additional mechanisms used by this pathogen when it infects humans. Precisely, we used the well-established model amoeba *Dictyostelium discoideum* as the predator. This species is a soil-derived amoeba that is widely used in host-pathogen interaction studies and is genetically tractable. Also central for our line of investigation, *D. discoideum* was used in initial studies that led to the discovery of the T6SS of *V. cholerae* (35). Indeed, Pukatzki and colleagues showed that the T6SS is used by toxigenic (but non-pandemic) *V. cholerae* to intoxicate *D. discoideum* and macrophages, while the ability of this nanomachine to also serve as a powerful killing device against other bacteria was only demonstrated in later years (36). This study and follow-up work done by others and us (29, 31) successfully demonstrated that *D. discoideum* poses a true predatory threat on *V. cholerae*. Of note, most of these investigations were based on plaque assays while visual inspections of the interaction between these organisms is still lacking. This work therefore aimed at establishing a microscopy-based protocol to visualize the interplay between bacteria and amoebae, initially focusing on the T6SS. Furthermore, we used this setup to investigate other possible virulence factors that might play important roles in the interaction of *V. cholerae* and eukaryotic grazers.

2.2 RESULTS AND DISCUSSION

2.2.1 *D. discoideum* amoebae keep normal morphology when co-cultured with *Klebsiella*

To visualize possible strategies used by *V. cholerae* to kill or avoid grazing by amoebae, first we had to establish and extensively optimize the co-culture conditions. By doing so, we arrived at a seamless time-lapse microscopy-based protocol that allowed coherent and reproducible visualization of the interactions between *D. discoideum* and *V. cholerae*. *D. discoideum* (10^6 cells) were seeded in rich medium into a µ-Dish (low wall 35mm) and left for adherence for at least 2 hours before being washed with SorC phosphate buffer. In the meantime, bacteria were grown aerobically, washed with PBS (Phosphate-Buffered Saline, pH 7.4), and resuspended in a OD_{600}-adjusted volume of SorC buffer. Finally, the bacteria were added to the amoebae at a multiplicity of infection (MOI) of 500 and the co-culture was imaged by confocal time-lapse microscopy (protocol details can be found in the Material and Methods section). SorC phosphate buffer was used as the co-culture medium as it lacks carbon sources and therefore avoids overgrowth of the medium by the planktonic bacteria. Given the lack of nutrients, amoebae rely on bacteria as their source of food.

To investigate potentially toxic *V. cholerae*-derived effects on the amoebae, we first had to establish control conditions in which the amoebae kept their active movement and ameboid morphology. For this reason, we first co-cultured *D. discoideum* with a non-encapsulated GFP-labelled *Klebsiella* strain (37) that is frequently used as a food source for *D. discoideum*. As expected, one hour after adding *Klebsiella*, *D. discoideum* displayed a normal amoeboid morphology (Figure 2.1a, second panel). Furthermore, as time passed the amoebae continued to graze on the *Klebsiella* cells, which was apparent by the decrease in GFP-labelled bacteria (Figure 2.1a, third and fourth panels).

In conclusion, we established a reliable time-lapse microscopy-based co-culture experimental setup to observe the interplay between *D. discoideum* and bacteria. Furthermore, these co-culture experiments with *Klebsiella* provided a comparative condition for normal cell morphology and amoebal behavior in general.

2.2.2 Exponentially growing *V. cholerae* kill *D. discoideum*

V. cholerae can employ its T6SS as a defense strategy against amoebae predation in solid media such as those used in plaque assays, where amoebal survival is scored by the number of grazing plaques formed on the bacterial lawn (29, 31, 35, 38). The live visualization of this protection mechanism has not been investigated by time lapse microscopy up to date.

After we established the appropriate experimental conditions (and controls) to observe the pathogen-host interaction by time-lapse confocal microscopy, we moved on to test the effect of *V. cholerae* on the amoebal population. We first tested the addition of overnight-grown *V. cholerae* (washed and OD_{600} adjusted) to the amoebae, as this condition seemed most comparable to the method used in solid plaque assays (35). Importantly and as done in previous work by others (35), we used a toxigenic non-pandemic *V. cholerae* strain that has a constitutively active T6SS (in our case, strain ATCC25872 (30) used throughout this study if not indicated otherwise), as this would allow us to witness T6SS-dependent phenotypes. As can be observed in Figure 2.1b (second panel), *D. discoideum* cells continued to display a normal cell morphology after one hour of co-incubation with *V. cholerae*. We also tested the amoebal reaction to bacteria that was back-diluted from the overnight culture and grown to early stationary phase (6 h), and likewise observed no signs of intoxication (Figure 2.1b, fourth panel).

Next, we co-incubated *D. discoideum* with *V. cholerae* that were grown to early exponential phase (2 h of growth after back dilution). Under these conditions, we witnessed a rapid intoxication of the amoebae (Figure 2.1b, third panel). Notably, we observed the same amoebal intoxication when they were exposed to an exponentially growing pandemic *V. cholerae* strain (A1552; Figure 2.1b, fifth panel). Pandemic strains are known to keep their T6SS silent under laboratory conditions, and utilize specific signals to activate their machinery (31, 39, 40). We therefore concluded that exponentially growing toxigenic *V. cholerae* are able to kill *D. discoideum* in liquid co-cultures and that this process is likely T6SS-independent.

2.2.3 Timeline of the amoebal intoxication phenotype

In order to better understand the chronology and steps of amoebal intoxication, we analyzed several time lapse experiments in which *D. discoideum* amoebae were co-incubated with exponentially growing *V. cholerae*. The phenotype consistently developed in a stepwise manner. Before co-incubation with bacteria, *D. discoideum* cells displayed normal ameboid morphology and moved extensively (Figure 2.1c, first panel). As soon as exponentially growing *V. cholerae* were added to the imaging device, the amoebae reacted strongly. Within few minutes after exposure to *V. cholerae,* the amoebae rounded up, detached from the bottom of the imaging device (Figure 2.1c, second panel), and progressively underwent blebbing at around 30 min post-primary contact (p.p.c.) (Figure 2.1c, third panel). The reaction proceeded with vacuolization of the already blebbing cells at around 60 min p.p.c. (Figure 2.1C, fourth panel) and resulted in their bursting at around 90-120 min p.p.c (Figure 2.1c, fifth panel). This cell death was unambiguously visible by their morphologically changed appearance and was further confirmed by loss of the GFP signal when GFP-tagged amoebal cell lines were used.

2.2.4 Amoebal intoxication is not driven by well-studied toxins and is contact-dependent

The next question we wanted to address was whether the phenotype relied on some of the well-studied toxins from *V. cholerae*, such as the multifunctional-autoprocessing repeats-in-toxin (MARTX), the pore-forming toxin hemolysin, the T6SS, or the haemagglutinin FrhA (explained below).

MARTX toxins are a variable family of very large proteins that are grouped together due to their type I secretion mode of transport and the presence of glycine-rich repeats that bind calcium, which is required for their proper folding and function (41). After secretion, these domain-structured toxins can bind to surface receptors on eukaryotic target cells, forming a pore that allows the residual domains to translocate across the membrane and exert their toxic effects intracellularly. Indeed, MARTX proteins often include toxic effector domains that disrupt the actin cytoskeleton of eukaryotic target cells (42). Notably, the MARTX toxin from *V. cholerae* (encoded by the *rtxA* gene) harbors such an actin-crosslinking domain (43). In a streptomycin-fed adult mice model of cholera, MARTX and specially hemolysin were demonstrated to be accessory toxins related to mouse lethality (44, 45). Hemolysin is highly conserved in *Vibrio* species (34) and is able to induce pore formation and vacuolation in eukaryotic target cells (46–49). Due to this pore formation, a proper regulation of hemolysin secretion is essential to allow *V. cholerae* to establish a reproductive niche within the contractile vacuole of aquatic amoeba *Acanthamoeba castellanii* (30). The flagellum-regulated hemagglutinin A (FrhA, named due to its motility-related regulation) contains MARTX-like features, such as a signature motif for type I secretion and Ca^{2+} binding sites. Furthermore, it also contains repetitive cadherin-like domains, which are commonly involved in calcium-dependent cell adhesion. Studies using classical and El Tor pandemic *V. cholerae* strains showed the involvement of FhrA in

phenotypes such as erythrocyte agglutination, binding to Hep-2 cells and chitin, early stages of biofilm formation, and mice colonization (50).

Upon testing the respective single mutants lacking these toxins/adhesin, toxicity against amoebae was still observed (data not shown). Furthermore, we had indirect evidence that the T6SS might not be involved in the amoebal intoxication phenotype (Figure 2.1B). The T6SS punctures eukaryotic target cells and delivers effectors such as the actin cross-linking domain (ACD), which intoxicates macrophages and grazing amoebae (31, 51, 52). We still considered, however, that there could be a compensatory or combinatory effect among all these proteins. We therefore genetically engineered a quadruple knockout in the *V. cholerae* ATCC25872 background that lacked *rtxA*, *hlyA*, *frhA,* and *vipA* (encoding a structural component of the T6SS). Surprisingly, despite their known anti-eukaryotic effects, the quadruple mutant was still toxic to amoebae (Figure 2.2a).

Next, we wondered whether the observed phenotype actually required direct contact between amoebae and bacteria. Answering this question could shed light onto the possible mode-of-action that was employed by the bacteria to bring about amoebal death. To test the contact dependency, we seeded amoebae on the bottom of a 24-well plate (appropriate for confocal microscopy), while adding *V. cholerae* inside a transwell insert that contained a 0.4 µm pore membrane at its bottom (Figure 2.2b). This setup allowed us to co-culture the amoebae and bacteria in a spatially separated manner, while still sharing the same culture buffer. Specifically, the membrane pores allow passage of potentially toxic molecules that are secreted by the bacteria. On the other hand, bacteria themselves would not be able to pass through the pores due to their size (roughly 1-2 µm) and would therefore not be able to harm the amoebae in case the intoxication was contact dependent.

As can be seen in Figure 2.2d, amoebae that were exposed to *V. cholerae* without contact continued to display a healthy cell morphology, comparable to their appearance before the addition of bacteria into the transwell insert (Figure 2.2c). A similar setup was used for the control condition in which case the transwell was not inserted thereby allowing bacteria-amoebae contact. As expected, under those conditions the amoebae were strongly impacted by the added bacteria (Figure 2.2e). We therefore conclude that the amoebal killing requires direct contact between bacteria and amoebae.

2.2.5 Toxicity does not require bacteria engulfment nor autophagy or the ESCRT machineries

Given that we gathered considerable information about the bacterial side of the phenotype, we decided to investigate the involvement of important host pathways. One unanswered question was whether the bacteria had to be phagocytosed in order to intoxicate the amoebae. Or, in other words, whether the defense mechanism put in place by *V. cholerae* was accomplished from outside the amoebal cell or inside the phagosomes (5). Phagocytosis is the process that amoebae use for feeding. Particles (larger than 200 nm, such as bacteria) are thereby recognized by surface receptors, leading to actin polymerization to distort the membrane around the particle, forming the phagocytic cup. After closure, the assembled phagosome undergoes a series of maturation steps that finally lead to degradation of the ingested particles (53). To check whether *V. cholerae* engulfment was necessary, we interrupted phagocytosis by incubating the amoebae with Cytochalasin D. This fungal toxin inhibits actin polymerization, therefore obstructing bacteria engulfment into phagosomes. Using this phagocytosis inhibitor, we observed, however, that exponentially growing *V. cholerae* were still able to efficiently intoxicate such non-phagocytic *D. discoideum* (Figure 2.3a; left panel). To support this finding, we also tested a genetically

engineered phagocytosis-impaired amoebal cell line (a triple vacuolin deletion strain, Δ*vac*ABC; kind gift from Prof. T. Soldati, together with the other amoebal cell lines mentioned below) (54). If bacteria engulfment were necessary, we would expect that these amoebae would be less intoxicated by *V. cholerae*, even at the considerably high MOI of 500 that we used in our experiments. However, as shown in Figure 2.3b (right panel), these phagocytosis-defective mutants were killed as efficiently as the parental wild-type amoebae, indicating that the defense mechanism is most likely put in place by the bacteria from outside of the eukaryotic cell.

The Endosomal Sorting Complexes Required for Transport (ESCRT) machinery is responsible for many functions inside the cell, including membrane damage repair (55, 56). When the plasma membrane is disrupted by small wounds caused by detergents or pore-forming toxins, there is an influx of Ca^{2+} through the wounds, resulting in rapid (around 30 seconds) recruitment of the initial components of the ESCRT complex to start membrane repair. Interestingly, ESCRT and autophagy were shown to work together to repair the *Mycobacterium*-containing vacuole, from which the bacterium tries to escape using its ESX-1 secretion system (57). We reasoned that if these machineries were necessary to attempt to repair the damage caused by *V. cholerae*, amoeba lines defective in their assembly would be hypersensitive to the bacterial killing process. We therefore tested amoeba lines lacking key components of either the ESCRT-I (Δ*tsg101*) (57) or autophagy (Δ*atg1)* (57) pathways, but observed that they were killed at comparable levels as the wild-type (Figure 2.3c, second and third panels). Furthermore, the double autophagy/ESCRT-I knockout (Δ*atg1*Δ*tsg101*) (57), which is very susceptible to membrane damage, was quickly killed by *V. cholerae* (Figure 2.3c, fourth panel). Future investigations will address the killing dynamics in more detail to better understand the involvement of repair mechanisms in a potential partial protection at lower MOIs.

Furthermore, as the ESCRT machinery repairs membrane damage, we speculated that tracking ESCRT assembly during our time-lapse imaging could 1) confirm that *V. cholerae* inflicts membrane damage to the amoebae; and 2) help resolve the time point at which the amoebal membranes are injured (bringing insight into a potential cause for toxicity). For that, we utilized a *D. discoideum* strain (Ax2 Ka *vps4*-GFP) in which the ESCRT AAA-ATPase Vps4 is fused to GFP (56, 57). This amoebal line allows the direct visualization of Vps4 recruitment through the formation of GFP puncta in areas of membrane damage. Using this strain, we observed that Vps4-GFP puncta appeared during amoebael intoxication by *V. cholerae*, though at a later stage, suggesting membrane damage as a consequence rather than initial cause (Figure 2.3d). Indeed, the amoebal blebbing only started after the first Vps4-GFP puncta appeared (Figure 2.3d; observe GFP puncta marked with arrows, starting in third panel), which suggests that the surrounding buffer enters the cells and ultimately leads to their bursting. If the Vps4 puncta had appeared at the very beginning of the intoxication progression, one could argue that an initial membrane damage, caused by *V. cholerae*, initiated the intoxication. However, based on the observation that the membrane damage occurs rather late, we suggest that the initial intoxication does not rely on pore formation by any bacterial effector(s).

2.2.6 The amoebal intoxication ability is not conserved among *Vibrio* species

Considering that our previous approaches did not give a clear indication of what could be the mechanism employed by *V. cholerae* to intoxicate amoebae, we decided to implement a comparative approach. As described above, both *V. cholerae* strains tested initially against amoebae are patient isolates. The pandemic T6SS-silent strain A1552 (serogroup O1) is connected to the large Peruvian outbreak from the 1990s and is part of the 7th pandemic clade (25–27, 58). The T6SS-active strain ATCC25872 is toxigenic but non-pandemic

(serogroup O37 with similarity to 6th pandemic classical strains), as it caused only a small outbreak in Czechoslovakia in the 1960s (25, 59). The fact that both of these strains were able to intoxicate amoebae (Figure 2.1b) indicates that the O-antigen (which defines the serogroup) - and the T6SS - are most likely not involved in the toxicity phenotype.

Still, *V. cholerae* is a diverse species. While large scale genomic analyses have shown the near clonality of pandemic strains, environmental isolates carry a high genomic diversity (25–28, 31; see chapter 3 of this thesis). We therefore reasoned that testing different environmental isolates against amoebae could help determine the mode-of-action behind the toxicity phenotype. For that, we tested *V. cholerae* environmental strains that were isolated from coastal California in 2004 and classified into four clades (A-D) based on comparative genomic hybridization (60). Specifically, we tested one representative strain from each clade (31): W10G (clade A); SA5Y (clade B); SL6Y (clade C) and SA10G (clade D). As can be seen in Figure 2.4A (upper row), all strains were able to kill *D. discoideum* when grown to exponential phase. To our surprise, however, strains SA5Y and SL6Y also intoxicated the amoebae when grown to stationary phase (Figure 2.4a, bottom row), in contrast to what we had previously observed with toxigenic strains (Figure 2.1b) and also with W10G and SA10G strains. This finding suggests that SA5Y and SL6Y might differentially regulate certain genes, allowing these strains to intoxicate amoebae even when grown to high cell density. We performed RNA sequencing to investigate the expression profile of pandemic *V. cholerae* A1552 and a selection of environmental isolates at these different time points. However, the immense transcriptomic differences between 2 h and 6 h growth did not allow us to suggest causative genes for the observed phenotype. Further inspection of the transcriptomics data will therefore be required in the future, which should be based on additional exclusion criteria.

Next, we wanted to explore whether the caused phenotype was genus (*Vibrio*) or species (*V. cholerae*) specific. We therefore investigated the effect of two species from the *V. harveyi* clade: *V. alginolyticus* (strain 12G01) and *V. parahaemolyticus* (strain POR1) (kind gift from Prof. D. Salomon; (61)). Both bacterial species are known to cause gastroenteritis in humans, usually related to the consumption of contaminated seafood. While *V. alginolyticus* – the most divergent species of the clade – did not cause any toxic reaction in *D. discoideum*, incubation with *V. parahaemolyticus* mimicked the toxic phenotype observed with *V. cholerae*, leading to the death of the whole amoebal population (Figure 2.4b, first and third images). *V. fischeri* is a member of the basal Photobacterium clade in the Vibrionaceae phylogenetic tree (61). These bacteria are free-living but can also live symbiotically inside the light organ of squids (*Euprymna scolopes*) or fish. In the squid-*V. fischeri* symbiosis, the bacteria produce light upon reaching high cell density thereby helping the squid to camouflage against predators by counterillumination. Interestingly, when exposed to *V. fischeri* (strain ES114), *D. discoideum* amoebae showed an intermediate toxicity phenotype (Figure 2.4b, second image). While approximately half of the amoebal population displayed the typical *V. cholerae*-like intoxication phenotype, the rest of the amoebae recovered from the initial “stress” and eventually resumed their movement.

Finally, we verified that the observed phenotype was not a general reaction to other bacteria. By using the same experimental setup described above, we tested an array of *E. coli* strains, including commensal species that were isolated from mice (62, 63), *B. subtilis,* and an encapsulated *K. pneumoniae* strain (different from the non-encapsulated strain used as a good food source control; Figure 2.1a). As shown in Figure 2.5, none of these bacteria prompted any type of toxic reaction/morphological changes in the amoebae. This was in

sharp contrast to the immediate rounding of the amoebae 5 min p.p.c. when in contact with *V. cholerae* (Figure 2.5, second panel in top row).

2.3 CONCLUSION AND FUTURE PERSPECTIVES

In this study, we developed a time-lapse microscopy-based approach to visualize and investigate the interactions between *D. discoideum* amoebae and *V. cholerae*. We observed that when exposed to exponentially growing toxigenic *V. cholerae*, amoebae display a toxicity response, with cells rapidly rounding up and detaching from the bottom of the imaging device. A well-synchronized and reproducible phenotype develops, where amoebae display blebbing and vacuolization features, finally lysing and dying within approximately 2 hours. Toxigenic *V. cholerae* grown to stationary phase, on the other hand, are innocuous to amoebae. Importantly, the pore-forming toxins MARTX and hemolysin, the adhesin FrhA and the T6SS from *V. cholerae* are not involved in the phenotype, as a quadruple mutant strain was as toxic to amoebae as the wild-type strain. Furthermore, when amoebae were exposed to *V. cholerae* but their contact was prevented, the cells remained healthy and kept their normal morphology. These findings suggest that amoebal intoxication requires contact with the bacteria.

Furthermore, we explored the potential involvement of certain host pathways in the observed phenotype. Phagocytosis was arrested with cytochalasin D or using a genetically engineered amoebal line (Δ*vacABC*) that is deficient for phagocytosis. Non-phagocytic amoebae were nevertheless promptly killed by *V. cholerae* in either of those conditions, suggesting that bacterial engulfment is not required for intoxication. Moreover, amoebae can combine the ESCRT and autophagy machineries to repair membrane damage. We therefore investigated whether these processes could be put in place by amoebae in an attempt to

rebuild damaged membranes during intoxication by *V. cholerae*. However, we observed that amoeba lines defective in the assembly of ESCRT-I (Δ*tsg101*) or autophagy (Δ*atg1)*, or a knockout for both pathways (Δ*atg1* Δ*tsg101*), were still killed by *V. cholerae* at comparable levels as wild-type amoebae. Moreover, an amoeba line in which the ESCRT AAA-ATPase Vps4 is fused to GFP was used to track the recruitment of ESCRT machineries to membrane damage sites when amoebae were exposed to *V. cholerae*. This allowed us to appreciate that membrane damage was inflicted rather late during the toxicity development, precisely around the blebbing stage. This reinforced our finding that intoxication does not rely on pore-forming toxins. Future work with these different amoeba lines will address specific aspects of the killing process and the possible involvement of cell repair mechanisms at lower MOIs.

Furthermore, *V. cholerae* as a species includes a wide diversity of strains, so we decided to employ a comparative approach and assess the ability of some environmental isolates to kill amoebae. These strains were isolated from a cholera-free region in coastal California in 2004 (60). Despite their varied genetic architecture (see chapter 3 of this thesis), all of these environmental isolates were able to intoxicate amoebae when grown to exponential phase. This result indicates that the observed phenotype does not require genomic elements that are exclusive to pathogenic lines (e.g., cholera toxin prophage, *Vibrio* pathogenicity islands 1 and 2). More importantly, two environmental isolates (SA5Y and SL6Y) were also capable of intoxicating amoebae when grown to stationary phase. This result indicates that these strains might regulate the causative gene(s) differently when compared to strains that are toxic only when exponentially growing. We went a step further in our comparative approach and tested whether the amoebae intoxication ability was also conserved among *Vibrio* species and other bacteria. While *E. coli*, *B. subtilis* and *K. pneumoniae* strains were harmless to amoebae, *V. parahaemolyticus* was as toxic as *V. cholerae*, hinting that the causative gene(s) is most likely *Vibrio* specific.

The next steps in this study will take advantage of these differences observed among *V. cholerae* strains. First, we will create a library of hybrid strains on a pandemic *V. cholerae* background. The clones of this library will contain different portions of the genome of an environmental strain that is toxic at stationary phase (SA5Y or SL6Y). Next, we will establish a medium-throughput alternative protocol for the co-incubation/microscopy experiments (for instance, by using GFP-labeled amoebae, which, upon lysis, lose their fluorescence signal). This protocol will be used to test the ability of these hybrid clones to intoxicate amoebae at stationary phase.

Furthermore, we have employed RNAseq to investigate the transcriptome of exponentially growing and stationary-phase pandemic and environmental strains of *V. cholerae*. Nevertheless, the differences in their expression profile due to growth phase alone are enormous, making it challenging to identify genes responsible for amoebae intoxication. Future sorting of the data using additional exclusion parameters is therefore required. We expect that the combination of in-depth RNAseq analyses with the hybrid strains approach will allow us to identify the mechanism responsible for amoebae killing.

2.4 MATERIAL AND METHODS

2.4.1 Bacterial and amoebal strains, and growth conditions

The bacterial and amoebal strains used in this study are listed in Supplementary table 2.1. Unless otherwise stated, all bacterial strains were grown aerobically in Lysogeny broth (LB; 10 g/ L of tryptone, 5 g/ L of yeast extract, 10 g/ L of sodium chloride; Carl Roth) or on LB agar plates at 30 °C. *V. fischeri* was grown in LB salt (LBS) medium (10 g/ L of tryptone, 5 g/ L of yeast extract, 20 g/ L of sodium chloride, 20 mM Tris-HCl pH 7.5 and 0.2% glycerol) (64) at 30 °C. Half-concentrated defined artificial seawater medium (0.5×DASW) containing

HEPES (4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid; Sigma) and vitamins (65) was used when strains were grown on chitin for natural transformation-based strain construction (see below).

D. discoideum amoebae lines were cultured in HL5 medium supplemented with glucose (5 g/ L peptone, 5 g/ L yeast extract, 5 g/ L tryptone, 1.2 g/ L KH_2PO_4, 0.35 g/ L Na_2HPO_4 and 10 g/ L glucose; Foremedium, UK). Sörensen's buffer (8 g /4 L of KH_2PO_4, 1.16 g /4 L of Na_2HPO_4 ; pH 6) supplemented with 50 µM $CaCl_2$ (66) was employed as the final resuspension buffer for bacteria and amoebae and also as the co-incubation buffer in microscopy experiments. All *D. discoideum* cell lines were a kind gift from the laboratory of Prof. Thierry Soldati (University of Geneva, Switzerland).

The following antibiotics were added, if required, at the given concentrations: kanamycin (75 µg/ mL), gentamycin (100 µg/ mL) and ampicillin (100 µg/ mL) for bacteria; G418 (5 µg/ mL) for amoebae.

2.4.2 Genetic engineering

V. cholerae strains were genetically engineered using the previously described TransFLP method (67–70). The technique depends on natural transformation activated by growth on chitin followed by addition of a PCR fragment that carries the desired genetic change. PCR amplifications were performed using GoTaq (Promega), Pwo (Roche) and Expand High Fidelity (Roche) polymerases according to the suppliers' recommendations. After screening by PCR using bacterial colonies as templates, genetically modified loci were checked by Sanger sequencing (Microsynth, Switzerland).

A variant of the mini-Tn7 transposon carrying an optimized version of constitutively expressed dsRed (*dsRed.T3[DNT]*) (39, 71) was integrated into the *V. cholerae*

chromosome via triparental mating (72) or by natural transformation on chitin flakes (for strain SA5Y).

2.4.3 Co-incubation experiments and confocal laser microscopy

Co-culture experiments to observe amoebae reactions to different bacteria were established in a 1mL final volume. Non-confluent amoebae cells were diluted into fresh culture medium (HL5c) at a concentration of 10^6 amoebae/ mL and seeded into a µ-Dish (low wall 35mm ibiTreat devices; 80136-IBI, Vitaris, Switzerland). After at least two hours of static incubation at 24 °C to allow proper adherence, amoebae were washed three times with SorC (i.e., the co-culture buffer). After the last wash, amoebae were let to adhere and adapt to the new buffer (770 µL volume) for 30 minutes. When analyzing the role of phagocytosis in the amoebal intoxication phenotype, 10 µM of Cytochalasin D (Thermo Fischer Scientific) was used to arrest actin polymerization.

Bacterial overnight cultures were used without (overnight cultures) or with (grown to exponential or stationary phase) 50-fold back dilution into fresh LB/LBS medium and growth at 30 °C and 180 rpm to different growth phases (exponential phase: 2 h; stationary phase: 6 h). After the appropriate growth time, cultures were harvested by centrifugation and the cell pellet was washed twice with PBS (Phosphate-Buffered Saline, pH 7.4, Life Technologies) and then resuspended in SorC buffer to each an OD_{600} of 3. As the amoebal reaction to *V. cholerae* is very fast, the microscopy set up needed to be prepared prior to bacteria being added to the dish. Therefore, the µ-Dish containing washed and adapted amoebae cells was positioned in the Zeiss LSM 700 inverted microscope (Zeiss, Switzerland) and checked before adding bacteria. 330 µL of bacteria (resulting in a MOI of 500) was added to the dish and the time lapse imaging was immediately started. The

amoebal population was imaged every 2 minutes (with Z-stack, as detached cells will go out of the initial focal plane). For transwell experiments, tissue culture inserts with 0.4 µm pores (Sarstedt, USA) were used in 24-well plates suitable for confocal microscopy (µ-Plate 24 well black ibiTreat; 82406, Vitaris, Switzerland). Confocal lasers scanning microscope (CLSM) imaging was used for almost all experiments. For few experiments in which a binocular microscope was used instead, the set up was adjusted to single wells of 12-well plates.

2.5 FIGURES

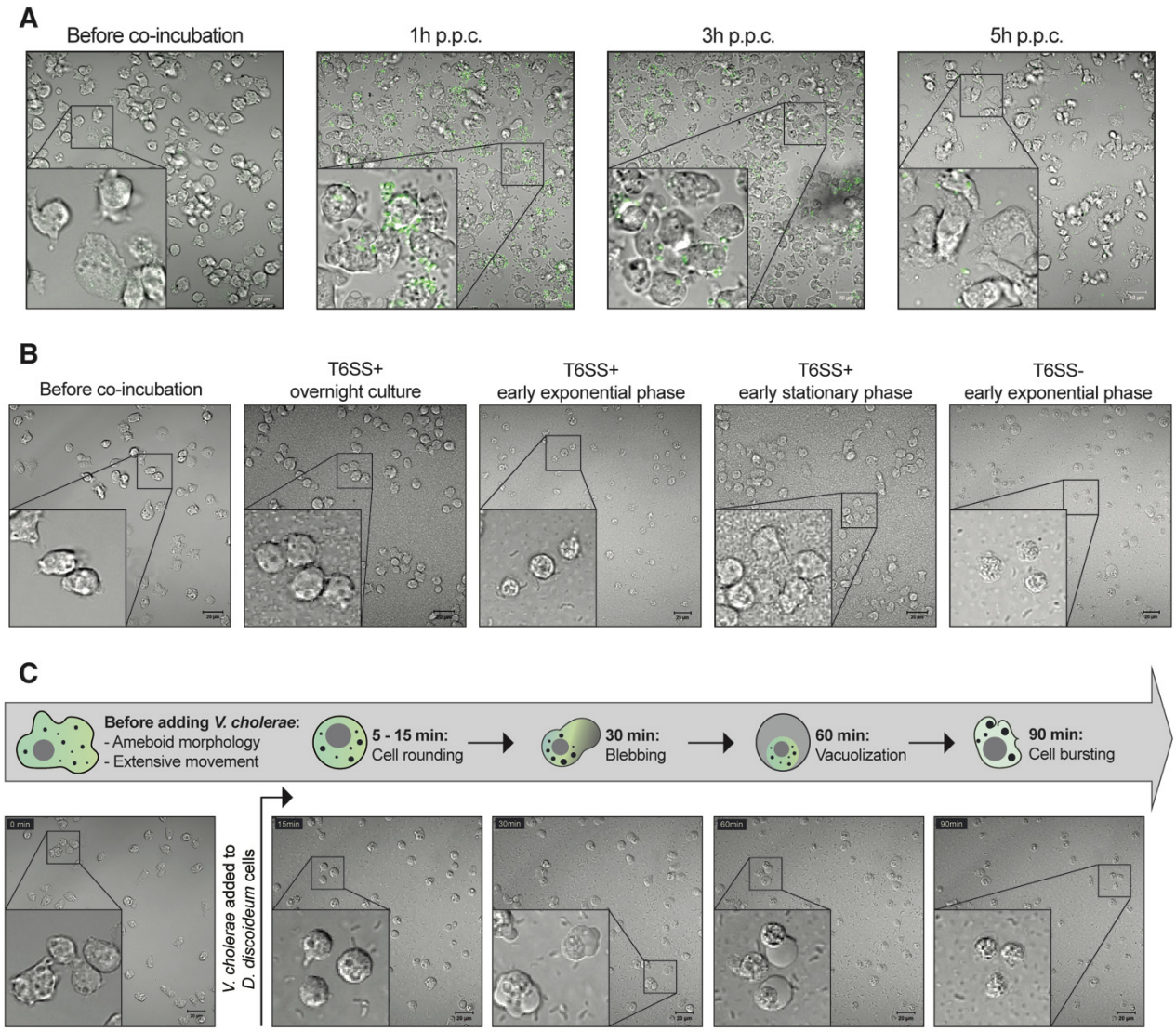


Figure 2.1 Morphological changes of *D. discoideum* after exposure to toxigenic *V. cholerae.*

(A) *D. discoideum* cells were mixed with non-toxic (e.g., non-encapsulated) GFP-labelled *Klebsiella* as food source. Images were taken before co-incubation started and 1, 3 and 5 h post primary contact (p.p.c.). Bacterial numbers decreased over time. **(B)** *D. discoideum* cells were exposed to *V. cholerae* (T6SS+ or T6SS-) grown to different growth phases (overnight, early exponential or early stationary phases). **(C)** Timeline illustration of the amoebal intoxication phenotype displayed after exposure to exponentially growing *V. cholerae*. Step I: before co-incubation started; amoebae with normal morphology and moving extensively; II (5-15 min): rounding up of amoebal cells and detachment from dish surface; III (30 min): cell blebbing; IV (60 min): vacuolization; V (90 min): cell burst. Lower left: magnification of the boxed area.

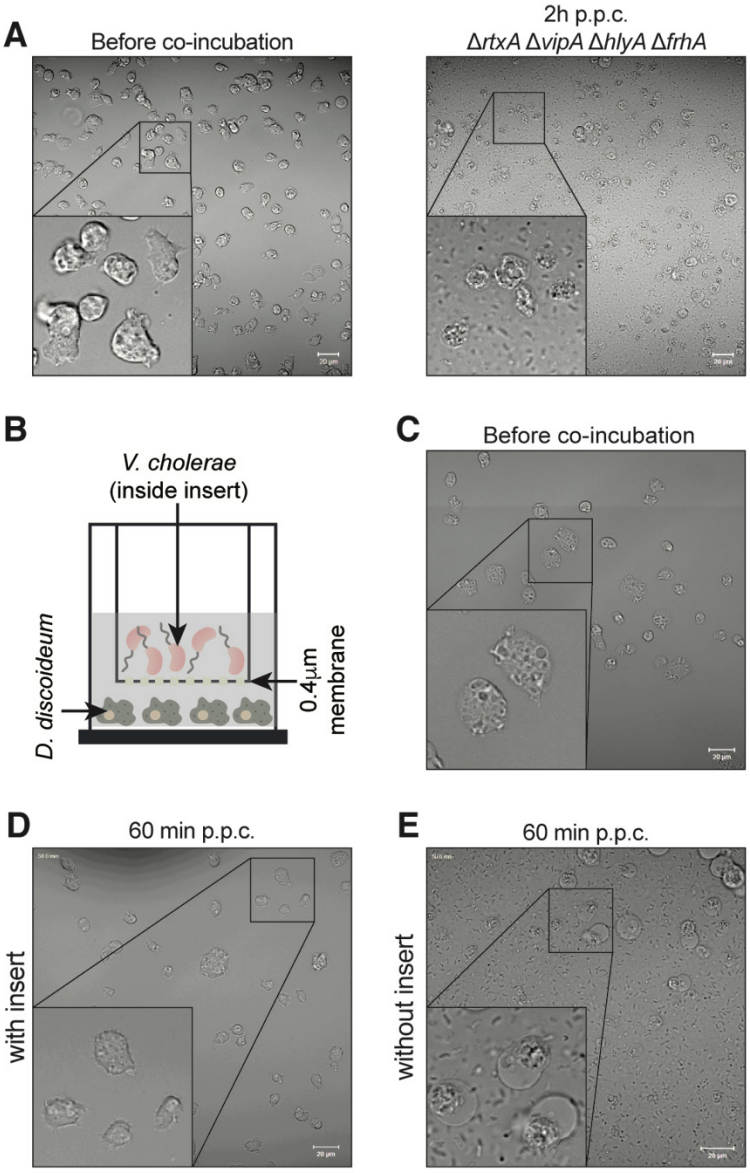


Figure 2.2 Amoebal intoxication does not depend on well-studied pore forming toxins or adhesins but requires direct contact.

(A) A toxigenic *V. cholerae* strain lacking the genes encoding the pore-forming toxin MARTX (*rtxA*), the type VI secretion system (T6SS), the hemolysin (*hlyA*) and the flagellum-regulated hemagglutinin A (*frhA*), was co-incubated with *D. discoideum*. Picture depicts the amoebal morphology 2 h p.p.c. **(B-E)** Transwell experiments as depicted in the scheme **(B)** to evaluate contact requirement for amoebal intoxication. **(C)** Amoebae before co-incubation. Amoeba cells were tested with (**D**) or without (**E**) inserts while exposed to *V. cholerae*. Lower left: magnification of the boxed area.

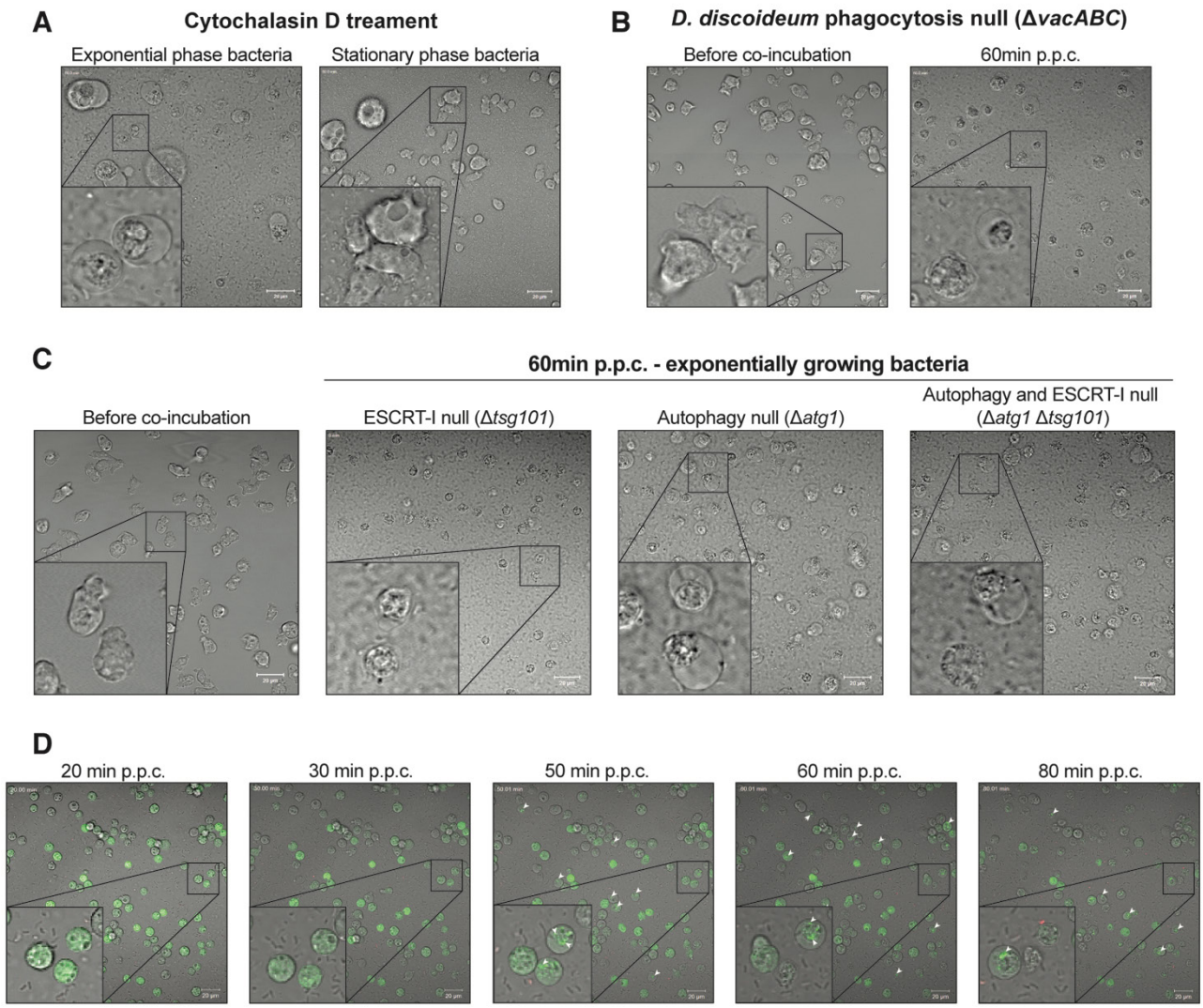


Figure 2.3 Amoebal intoxication is not linked to phagocytosis, autophagy, or the ESCRT pathways.

(A) *D. discoideum* amoebae were treated with cytochalasin D to inhibit actin polymerization and therefore phagocytosis before co-incubation with *V. cholerae* bacteria pre-grown to exponential or stationary phase. Pictures were taken 60 min p.p.c. Intoxication of a genetically engineered *D. discoideum* mutants defective for phagocytosis (Δ*vacABC;* **B**), the ESCRT-I pathway (Δ*tsg101;* **C**), autophagy (Δ*atg1;* **C**), or a double amoebal mutant (Δ*atg1* Δ*tsg101;* **C**). Morphological changes were comparable to the intoxicated WT amoebae. Pictures before addition of bacteria were included as controls. (**D**) *Vibrio*-caused intoxication leads to membrane damage. GFP puncta were observed when *D. discoideum* containing the AAA-ATPase *Vsp4* fused to GFP were imaged, hinting at assembled ESCRT machineries. Lower left: magnification of the boxed area.

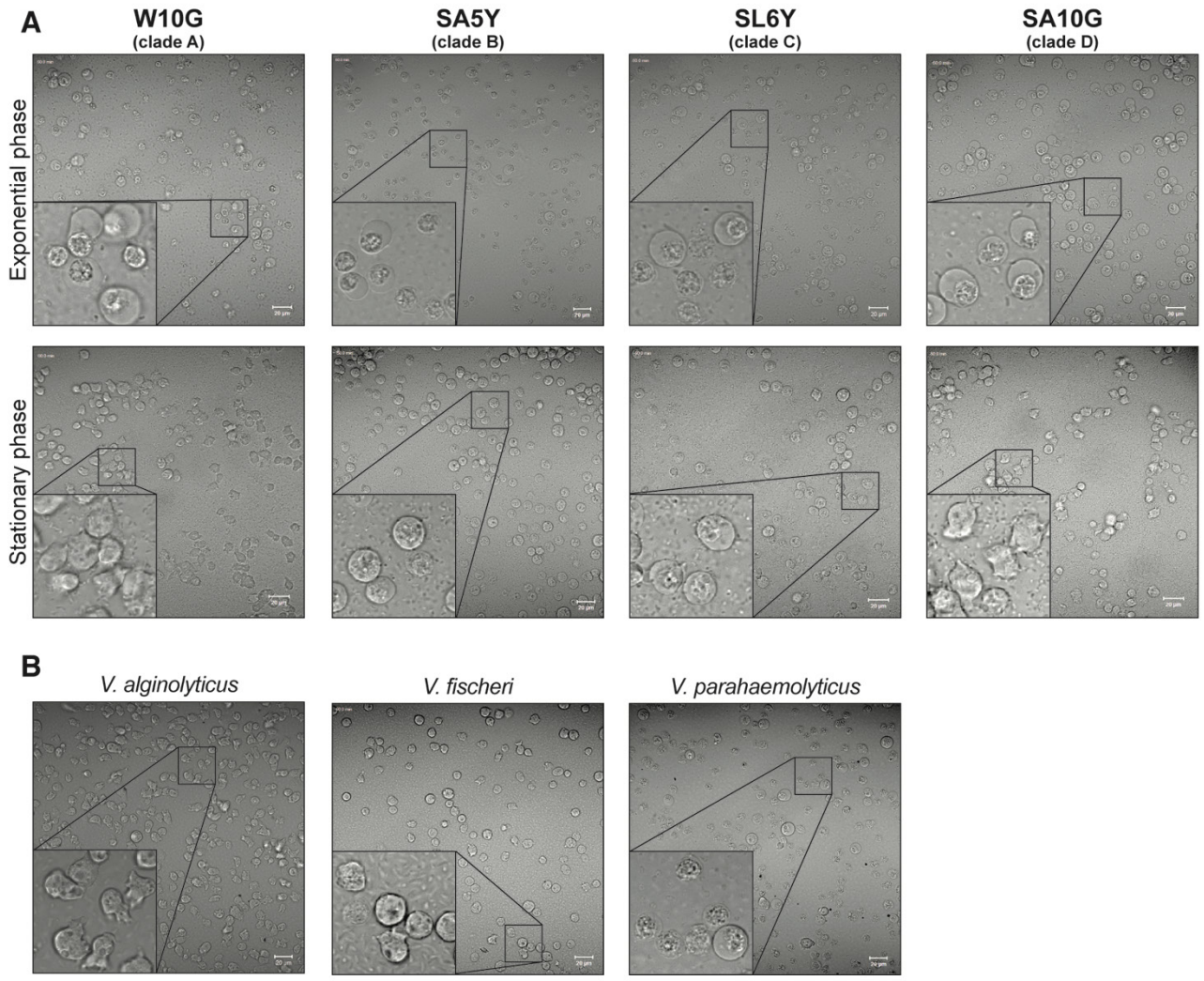


Figure 2.4 Not all *Vibrio* strains intoxicate amoebae.

D. discoideum amoebae were co-incubated with different *V. cholerae* strains (**A**) or *Vibrio* species (**B**) and imaged 60 min p.p.c. (**A)** Four *V. cholerae* strains from a collection of environmental isolates were tested for their ability to intoxicate *D. discoideum* after their growth to exponential (upper row) or stationary phase (lower row). Note the unexpected toxicity of strains SA5Y and SL6Y when grown to stationary phase. (**B**) Different phenotypes were observed upon amoebal incubation with *V. alginolyticus* (no morphological changes), *V. fischeri* (intermediate toxicity) or *V. parahaemolyticus* (full toxicity). Lower left: magnification of the boxed area.

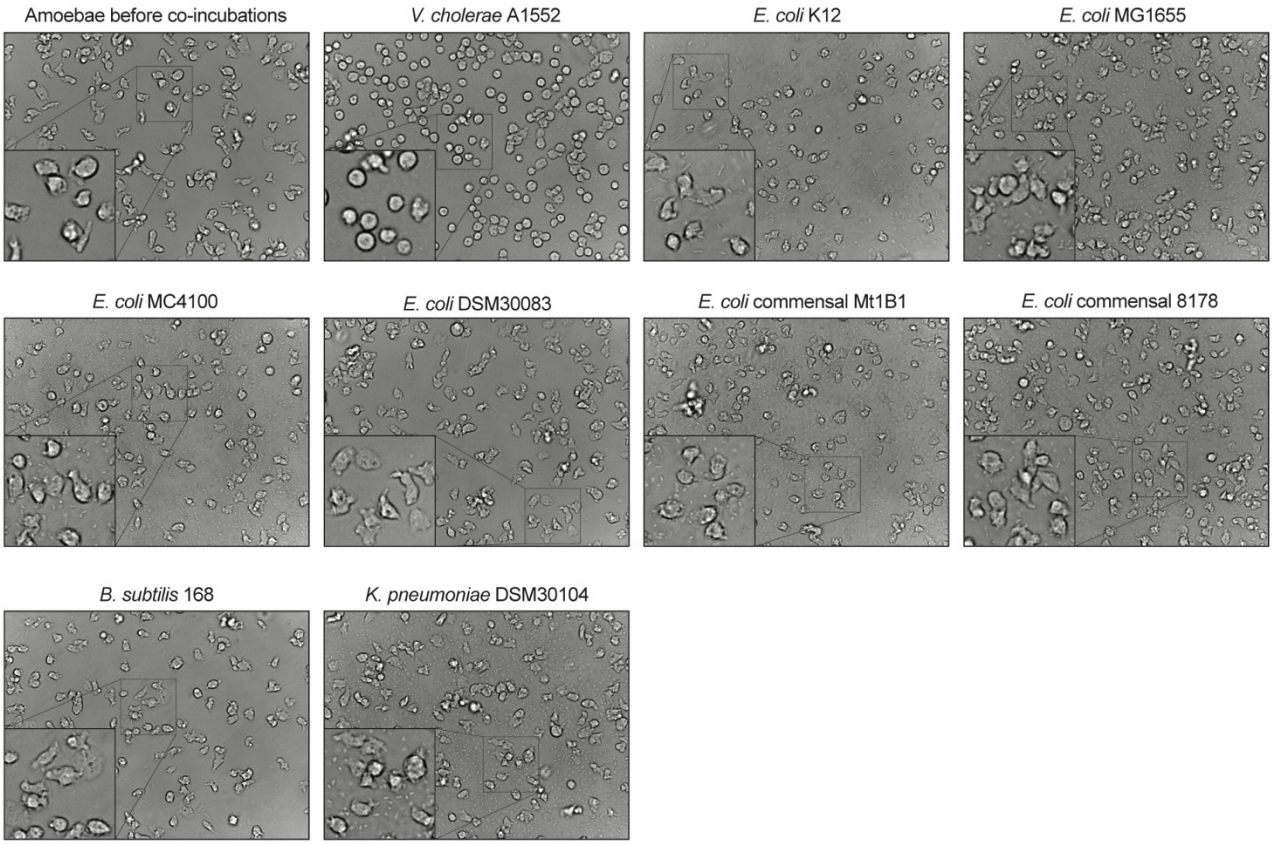


Figure 2.5 Amoebal intoxication is not a reaction to all bacteria.

D. discoideum amoebae kept their normal morphology after being exposed to exponentially growing strains of *E. coli*, *K. pneumoniae*, and *B. subtilis*. Pre-inoculated amoebae (first panel) and those incubated with intoxicating *V. cholerae* (second panel) are shown for comparison.

2.6 SUPPLEMENTARY TABLES

Supplementary table 2.1 *V. cholerae*, *E. coli*, *K. pneumoniae, B. subtilis and D. discoideum* strains used in this study.

Strain names	Genotype / description*	Internal strain number	Reference
V. cholerae strains			
A1552	Wild-type, O1 El Tor Inaba; Rif^R	MB_1	(58)
A1552-dsRed	A1552 with mTn7-dsRed.T3[DNT]; Rif^R, Gent^R	MB_1524	(39)
ATCC25872	*V. cholerae* non-O1 strain (O37); isolated in 1965, Czechoslovakia; intermediate resistant to Strep	MB_276	(59)
ATCC25872-dsRed	ATCC25872 with mTn7-dsRed.T3[DNT]; intermediate resistant to Strep, Gent^R	MB_6023	This study
ATCC25872Δ*rtxA*, Δ*vipA*, Δ*hlyA*, Δ*fhrA*	ATCC25872 deleted for *rtxA*, *vipA* and *hlyA* (TransFLP) and with *aph* insertion in *fhrA* (VC1620); intermediate resistant to Strep, Kan^R	ND_139	This study
W10G	Environmental isolate (clade A) collected in Waddell Creek (CA, USA) in October 2004	MB_5537	(60)
SA5Y	Environmental isolate (clade B) collected in Old Salinas River (CA, USA) in May 2004	MB_353	(60)
SA5Y-dsRed	SA5Y with mTn7-dsRed.T3[DNT]; Gent^R	ND_89	This study
SL6Y	Environmental isolate (clade C) collected in San Lorenzo River (CA, USA) in June 2004	MB_953	(60)
SA10G	Environmental isolate (clade D) collected in Old Salinas River (CA, USA) in October 2004	MB_5539	(60)
Non-cholera *Vibrio* species			
***V. parahaemolyticus* POR1**	Wild-type; RIMD 2210633 Δ*tdhAS*	MB_5862	(73)
***V. alginolyticus* 12G01**	Wild-type	MB_5857	(74)
***V. fischeri* ES114**	Wild-type	MB_5869	ATCC-700601; via LGC standards
E. coli, K. pneumoniae and *B. subtilis* strains			
K. pneumoniae	*Klebsiella* sp. / pANT5 (GFP); Amp^R	MB_6380	(75)
***K. pneumoniae* DSM30104**	Wild-type; ATCC 13883; Capsular serovar: 3; Biosafety class 2	MB_4327	German Collection of Microorganisms and Cell Cultures (DSMZ)
***E. coli* K12**		MB_2903	laboratory collection
***E. coli* MG1655**		MB_2904	laboratory collection
***E. coli* MC4100**	Strep^R	MB_2905	laboratory collection
***E. coli* DSM30083**	Wild-type; ATCC 1177; Serovar O1:K1:H7; Biosafety class 2	MB_4326	DSMZ
***E. coli* Mt1B1**	Mouse commensal isolate; Cm^R	MB_4992	(62)

***E. coli* 8178**	Mouse commensal isolate; Cm^R	MB_4993	(63)
***B. subtilis* 168**		MB_3230	laboratory collection
D. discoideum strains			
Ax2(Ka)	Wild-type	MB_6279	laboratory collection**
Ax2(Ka) pDM323 *vps4*-GFP	*vps4* cDNA (DDB_G0284347) in pDM323 plasmid; $G418^R$	MB_6282	(57)
Ax2(Ka) ABD-GFP	Actin binding domain-GFP; $G418^R$	MB_6281	laboratory collection**
Ax2(Ka) Δ*tsg101*	ESCRT-I null	MB_7918	(57)
Ax2(Ka) Δ*atg1*	Autophagy null	MB_7919	(76)
Ax2(Ka) Δ*atg1* Δ*tsg101*	Autophagy and ESCRT-I null	MB_7920	(57)
Ax2(Ka) Δ*vacABC*	Phagocytosis null	MB_7921	(54)

* locus tags belong to reference strain N16961 according to (77).

**received from laboratory collection of Prof. Thierry Soldati, University of Geneva.

2.7 REFERENCES

1. Brüssow,H. (2007) Bacteria between protists and phages: from antipredation strategies to the evolution of pathogenicity. *Mol. Microbiol.*, **65**, 583–589.

2. Flemming,H.C. and Wuertz,S. (2019) Bacteria and archaea on Earth and their abundance in biofilms. *Nat. Rev. Microbiol.*, **17**, 247–260.

3. Granato,E.T., Meiller-Legrand,T.A. and Foster,K.R. (2019) The evolution and ecology of bacterial warfare. *Curr. Biol.*, **29**, R521–R537.

4. Peterson,S.B., Bertolli,S.K. and Mougous,J.D. (2020) The central role of interbacterial antagonism in bacterial fife. *Curr. Biol.*, **30**, R1203–R1214.

5. Matz,C. and Kjelleberg,S. (2005) Off the hook - How bacteria survive protozoan grazing. *Trends Microbiol.*, **13**, 302–307.

6. Gao,L., Altae-Tran,H., Böhning,F., Makarova,K.S., Segel,M., Schmid-Burgk,J.L., Koob,J., Wolf,Y.I., Koonin,E. V. and Zhang,F. (2020) Diverse enzymatic activities mediate antiviral immunity in prokaryotes. *Science*, **369**, 1077–1084.

7. Doron,S., Melamed,S., Ofir,G., Leavitt,A., Lopatina,A., Keren,M., Amitai,G. and Sorek,R. (2018) Systematic discovery of antiphage defense systems in the microbial pangenome. *Science*, **359**, eaar4120.

8. Justice,S.S., Hunstad,D.A., Cegelski,L. and Hultgren,S.J. (2008) Morphological plasticity as a bacterial survival strategy. *Nat. Rev. Microbiol.*, **6**, 162–168.

9. Pernthaler,J. (2005) Predation on prokaryotes in the water column and its ecological implications. *Nat. Rev. Microbiol.*, **3**, 537–546.

10. Casadevall,A. (2008) Evolution of intracellular pathogens. *Annu. Rev. Microbiol.*, **62**, 19–33.

11. Scheid,P. (2014) Relevance of free-living amoebae as hosts for phylogenetically diverse microorganisms. *Parasitol. Res.*, **113**, 2407–2414.

12. Strassmann,J.E. and Shu,L. (2017) Ancient bacteria–amoeba relationships and pathogenic animal bacteria. *PLoS Biol.*, **15**, e2002460.

13. Brown,M. and Barker,J. (1999) Unexplored reservoirs of pathogenic bacteria: protozoa and biofilms. *Trends Microbiol.*, **7**, 46–50.

14. Greub,G. and Raoult,D. (2004) Microorganisms resistant to free-living amoebae. *Clin. Microbiol. Rev.*, **17**, 413–433.

15. Molmeret,M., Horn,M., Wagner,M., Santic,M. and Kwaik,Y.A. (2005) Amoebae as training grounds for intracellular bacterial pathogens. *Appl. Environ. Microbiol.*, **71**, 20–28.

16. Escoll,P., Mondino,S., Rolando,M. and Buchrieser,C. (2015) Targeting of host organelles by pathogenic bacteria: A sophisticated subversion strategy. *Nat. Rev. Microbiol.*, **14**, 5–19.

17. Hilbi,H., Weber,S.S., Ragaz,C., Nyfeler,Y. and Urwyler,S. (2007) Environmental predators as models for bacterial pathogenesis. *Environ. Microbiol.*, **9**, 563–575.

18. Adiba,S., Nizak,C., van Baalen,M., Denamur,E. and Depaulis,F. (2010) From grazing resistance to pathogenesis: The coincidental evolution of virulence factors. *PLoS One*, **5**, e11882.

19. Clemens,J.D., Nair,G.B., Ahmed,T., Qadri,F. and Holmgren,J. (2017) Cholera. *Lancet*, **390**, 1539–1549.

20. World Health Organization (WHO) (2006) Communicable diseases following natural disasters Risk assessment and priority interventions.

21. Faruque,S.M., Albert,M.J. and Mekalanos,J.J. (1998) Epidemiology, genetics, and ecology of toxigenic

Vibrio cholerae. *Microbiol. Mol. Biol. Rev.*, **62**, 1301–14.

22. Cottingham,K.L., Chiavelli,D.A. and Taylor,R.K. (2003) Environmental microbe and human pathogen: the ecology and microbiology of *Vibrio cholerae*. *Front Ecol Environm*, **1**, 80–86.

23. Taylor,R.K., Miller,V.L., Furlong,D.B. and Mekalanos,J.J. (1987) Use of phoA gene fusions to identify a pilus colonization factor coordinately regulated with cholera toxin. *Proc. Natl. Acad. Sci.*, **84**, 2833–2837.

24. Waldor,M.K. and Mekalanos,J.J. (1996) Lysogenic conversion by a filamentous phage encoding cholera toxin. *Science*, **272**, 1910–1914.

25. Chun,J., Grim,C.J., Hasan,N.A., Lee,J.H., Choi,S.Y., Haley,B.J., Taviani,E., Jeon,Y.-S., Kim,D.W., Lee,J.-H., *et al.* (2009) Comparative genomics reveals mechanism for short-term and long-term clonal transitions in pandemic *Vibrio cholerae*. *Proc. Natl. Acad. Sci.*, **106**, 15442–15447.

26. Mutreja,A., Kim,D.W., Thomson,N.R., Connor,T.R., Lee,J.H., Kariuki,S., Croucher,N.J., Choi,S.Y., Harris,S.R., Lebens,M., *et al.* (2011) Evidence for several waves of global transmission in the seventh cholera pandemic. *Nature*, **477**, 462–465.

27. Domman,D., Quilici,M.-L., Dorman,M.J., Njamkepo,E., Mutreja,A., Mather,A.E., Delgado,G., Morales-Espinosa,R., Grimont,P.A.D., Lizarraga-Partida,M.L., *et al.* (2017) Integrated view of *Vibrio cholerae* in the Americas. *Science*, **358**, 789–793.

28. Weill,F.-X., Domman,D., Njamkepo,E., Tarr,C., Rauzier,J., Fawal,N., Keddy,K.H., Salje,H., Moore,S., Mukhopadhyay,A.K., *et al.* (2017) Genomic history of the seventh pandemic of cholera in Africa. *Science*, **358**, 785–789.

29. Unterweger,D., Kitaoka,M., Miyata,S.T., Bachmann,V., Brooks,T.M., Moloney,J., Sosa,O., Silva,D., Duran-Gonzalez,J., Provenzano,D., *et al.* (2012) Constitutive type VI secretion system expression gives *Vibrio cholerae* intra- and interspecific competitive advantages. *PLoS One*, **7**, e48320.

30. Van der Henst,C., Vanhove,A.S., Drebes Dörr,N.C., Stutzmann,S., Stoudmann,C., Clerc,S., Scrignari,T., Maclachlan,C., Knott,G. and Blokesch,M. (2018) Molecular insights into *Vibrio cholerae*'s intra-amoebal host-pathogen interactions. *Nat. Commun.*, **9**, 3460.

31. Drebes Dörr,N.C. and Blokesch,M. (2020) Interbacterial competition and anti-predatory behavior of environmental *Vibrio cholerae* strains. *Environ. Microbiol.*, **22**, 4485–4504.

32. Bernardy,E.E., Turnsek,M.A., Wilson,S.K., Tarr,C.L. and Hammer,B.K. (2016) Diversity of clinical and environmental isolates of *Vibrio cholerae* in natural tansformation and contact-dependent bacterial killing indicative of type VI secretion system activity. *Appl. Environ. Microbiol.*, **82**, 2833–2842.

33. Boyer,F., Fichant,G., Berthod,J., Vandenbrouck,Y. and 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**.

34. Zhang,X.H. and Austin,B. (2005) Haemolysins in *Vibrio* species. *J. Appl. Microbiol.*, **98**, 1011–1019.

35. Pukatzki,S., Ma,A.T., Sturtevant,D., Krastins,B., Sarracino,D., Nelson,W.C., Heidelberg,J.F. and Mekalanos,J.J. (2006) Identification of a conserved bacterial protein secretion system in *Vibrio cholerae* using the *Dictyostelium* host model system. *Proc. Natl. Acad. Sci.*, **103**, 1528–33.

36. MacIntyre,D.L., Miyata,S.T., Kitaoka,M. and Pukatzki,S. (2010) The *Vibrio cholerae* type VI secretion system displays antimicrobial properties. *Proc. Natl. Acad. Sci.*, **107**, 19520–19524.

37. Benghezal,M., Fauvarque,M.O., Tournebize,R., Froquet,R., Marchetti,A., Bergeret,E., Lardy,B., Klein,G., Sansonetti,P., Charette,S.J., *et al.* (2006) Specific host genes required for the killing of *Klebsiella* bacteria by phagocytes. *Cell. Microbiol.*, **8**, 139–148.

38. Miyata,S.T., Kitaoka,M., Brooks,T.M., McAuley,S.B. and Pukatzki,S. (2011) *Vibrio cholerae* requires the type VI secretion system virulence factor VasX to kill *Dictyostelium discoideum*. *Infect. Immun.*, **79**,

2941–2949.

39. Borgeaud,S., Metzger,L.C., Scrignari,T. and Blokesch,M. (2015) The type VI secretion system of *Vibrio cholerae* fosters horizontal gene transfer . *Science*, **347**, 63–67.

40. Metzger,L.C., Stutzmann,S., Scrignari,T., Van der Henst,C., Matthey,N. and Blokesch,M. (2016) Independent regulation of type VI secretion in *Vibrio cholerae* by TfoX and TfoY. *Cell Rep.*, **15**, 951–958.

41. Linhartová,I., Bumba,L., Mašn,J., Basler,M., Osička,R., Kamanová,J., Procházková,K., Adkins,I., HejnováHolubová,J., Sadílková,L., *et al.* (2010) RTX proteins: A highly diverse family secreted by a common mechanism. *FEMS Microbiol. Rev.*, **34**, 1076–1112.

42. Satchell,K.J.F. (2015) Multifunctional-autoprocessing toxins of Vibrios. *Microbiol. Spectr.*, **3**, VE-0002-2014.

43. Sheahan,K.-L., Cordero,C.L. and Fullner Satchell,K.J. (2004) Identification of a domain within the multifunctional *Vibrio cholerae* RTX toxin that covalently cross-links actin. *Proc. Natl. Acad. Sci.*, **101**, 9798–9803.

44. Olivier,V., Haines,G.K., Tan,Y. and Fullner Satchell,K.J. (2007) Hemolysin and the multifunctional autoprocessing RTX toxin are virulence factors during intestinal infection of mice with *Vibrio cholerae* El Tor O1 strains. *Infect. Immun.*, **75**, 5035–5042.

45. Olivier,V., Salzman,N.H. and Fullner Satchell,K.J. (2007) Prolonged colonization of mice by *Vibrio cholerae* El Tor O1 depends on accessory toxins. *Infect. Immun.*, **75**, 5043–5051.

46. Coelho,A., Andrade,J.R.C., Vicente,A.C.P. and Dirita,V.J. (2000) Cytotoxic cell vacuolating activity from *Vibrio cholerae* hemolysin. *Infect. Immun.*, **68**, 1700–1705.

47. Menzl,K., Maier,E., Chakraborty,T. and Benz,R. (1996) HlyA hemolysin of *Vibrio cholerae* O1 biotype El Tor: identification of the hemolytic complex and evidence for the formation of anion-selective ion-permeable channels. *Eur. J. Biochem.*, **240**, 646–654.

48. Figueroa-Arredondo,P., Heuser,J.E., Akopyants,N.S., Morisaki,J.H., Giono-Cerezo,S., Enríquez-Rincón,F. and Berg,D.E. (2001) Cell vacuolation caused by *Vibrio cholerae* hemolysin. *Infect. Immun.*, **69**, 1613–1624.

49. Peraro,M.D. and Van Der Goot,F.G. (2016) Pore-forming toxins: Ancient, but never really out of fashion. *Nat. Rev. Microbiol.*, **14**, 77–92.

50. Syed,K.A., Beyhan,S., Correa,N., Queen,J., Liu,J., Peng,F., Satchell,K.J.F., Yildiz,F. and Klose,K.E. (2009) The *Vibrio cholerae* flagellar regulatory hierarchy controls expression of virulence factors. *J. Bacteriol.*, **191**, 6555–6570.

51. Pukatzki,S., Ma,A.T., Revel,A.T., Sturtevant,D. and Mekalanos,J.J. (2007) Type VI secretion system translocates a phage tail spike-like protein into target cells where it cross-links actin. *Proc. Natl. Acad. Sci.*, **104**, 15508–15513.

52. Ma,A.T., McAuley,S., Pukatzki,S. and Mekalanos,J.J. (2009) Translocation of a *Vibrio cholerae* type VI secretion effector requires bacterial endocytosis by host cells. *Cell Host Microbe*, **5**, 234–243.

53. Dunn,J.D., Bosmani,C., Barisch,C., Raykov,L., Lefrançois,L.H., Cardenal-Muñoz,E., López-Jiménez,A.T. and Soldati,T. (2018) Eat prey, live: *Dictyostelium discoideum* as a model for cell-autonomous defenses. *Front. Immunol.*, **8**.

54. Bosmani,C., Leuba,F., Hanna,N., Bach,F., Burdet,F., Pagni,M., Hagedorn,M. and Soldati,T. (2020) Vacuolins and myosin VII are required for phagocytic uptake and phagosomal membrane recycling in *Dictyostelium discoideum*. *J. Cell Sci.*, **133**.

55. Vietri,M., Radulovic,M. and Stenmark,H. (2020) The many functions of ESCRTs. *Nat. Rev. Mol. Cell Biol.*, **21**, 25–42.

56. Schmidt,O. and Teis,D. (2012) The ESCRT machinery. *Curr. Biol.*, **22**, R116–R120.

57. López-Jiménez,A.T., Cardenal-Muñoz,E., Leuba,F., Gerstenmaier,L., Barisch,C., Hagedorn,M., King,J.S. and Soldati,T. (2018) The ESCRT and autophagy machineries cooperate to repair ESX-1-dependent damage at the *Mycobacterium*-containing vacuole but have opposite impact on containing the infection. *PLoS Pathog.*, **14**, e1007501.

58. Yildiz,F.H. and Schoolnik,G.K. (1998) Role of *rpoS* in stress survival and virulence of *Vibrio cholerae*. *J. Bacteriol.*, **180**, 773–784.

59. Aldova,E., Laznickova,K., Stepankova,E. and Lietava,J. (1968) Isolation of nonagglutinable vibrios from an enteritis outbreak in Czechoslovakia. *J. Infect. Dis.*, **118**, 25–31.

60. Keymer,D.P., Miller,M.C., Schoolnik,G.K. and Boehm,A.B. (2007) Genomic and phenotypic diversity of coastal *Vibrio cholerae* strains is linked to environmental factors. *Appl. Environ. Microbiol.*, **73**, 3705–3714.

61. Dikow,R.B. and Smith,W.L. (2013) Genome-level homology and phylogeny of Vibrionaceae (Gammaproteobacteria: Vibrionales) with three new complete genome sequences. *BMC Microbiol.*, **13**.

62. Brugiroux,S., Beutler,M., Pfann,C., Garzetti,D., Ruscheweyh,H.J., Ring,D., Diehl,M., Herp,S., Lötscher,Y., Hussain,S., *et al.* (2016) Genome-guided design of a defined mouse microbiota that confers colonization resistance against *Salmonella enterica* serovar Typhimurium. *Nat. Microbiol.*, **2**.

63. Stecher,B., Denzler,R., Maier,L., Bernet,F., Sanders,M.J., Pickard,D.J., Barthel,M., Westendorf,A.M., Krogfelt,K.A., Walker,A.W., *et al.* (2012) Gut inflammation can boost horizontal gene transfer between pathogenic and commensal Enterobacteriaceae. *Proc. Natl. Acad. Sci. U. S. A.*, **109**, 1269–1274.

64. Metzger,L.C., Matthey,N., Stoudmann,C., Collas,E.J. and Blokesch,M. (2019) Ecological implications of gene regulation by TfoX and TfoY among diverse Vibrio species. *Environ. Microbiol.*, **21**, 2231–2247.

65. Meibom,K.L., Blokesch,M., Dolganov,N.A., Wu,C.Y. and Schoolnik,G.K. (2005) Chitin induces natural competence in *Vibrio cholerae*. *Science*, **310**, 1824–1827.

66. Gerisch, G; Lüderitz, O; Ruschmann,E. (1967) Antikörper fördern die Phagozytose von Bakterien durch Amöben. *Zsch Naturforsch.*, **22b**, 109.

67. Blokesch,M. (2012) TransFLP - A method to genetically modify *Vibrio cholerae* based on natural transformation and FLP-recombination. *J. Vis. Exp.*, 10.3791/3761.

68. Borgeaud,S. and Blokesch,M. (2013) Overexpression of the tcp gene cluster using the T7 RNA polymerase/promoter system and natural transformation-mediated genetic engineering of *Vibrio cholerae*. *PLoS One*, **8**, e53952.

69. de Souza Silva,O. and Blokesch,M. (2010) Genetic manipulation of *Vibrio cholerae* by combining natural transformation with FLP recombination. *Plasmid*, **64**, 186–195.

70. Marvig,R.L. and Blokesch,M. (2010) Natural transformation of *Vibrio cholerae* as a tool-optimizing the procedure. *BMC Microbiol.*, **10**.

71. Dunn,A.K., Millikan,D.S., Adin,D.M., Bose,J.L. and Stabb,E. V. (2006) New rfp- and pES213-derived tools for analyzing symbiotic *Vibrio fischeri* reveal patterns of infection and *lux* expression *in situ*. *Appl. Environ. Microbiol.*, **72**, 802–810.

72. Bao,Y., Lies,D.P., Fu,H. and Roberts,G.P. (1991) An improved Tn7-based system for the single-copy insertion of cloned genes into chromosomes of gram-negative bacteria. *Gene*, **109**, 167–168.

73. Salomon,D., Gonzalez,H., Updegraff,B.L. and 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.

74. Salomon,D., Klimko,J.A., Trudgian,D.C., Kinch,L.N., Grishin,N. V., Mirzaei,H. and Orth,K. (2015) Type VI secretion system toxins horizontally shared between marine bacteria. *PLoS Pathog.*, **11**, e1005128.

75. Benghezal,M., Fauvarque,M.O., Tournebize,R., Froquet,R., Marchetti,A., Bergeret,E., Lardy,B., Klein,G., Sansonetti,P., Charette,S.J., *et al.* (2006) Specific host genes required for the killing of Klebsiella bacteria by phagocytes. *Cell. Microbiol.*, **8**, 139–148.

76. Cardenal-Muñoz,E., Arafah,S., López-Jiménez,A.T., Kicka,S., Falaise,A., Bach,F., Schaad,O., King,J.S., Hagedorn,M. and Soldati,T. (2017) *Mycobacterium marinum* antagonistically induces an autophagic response while repressing the autophagic flux in a TORC1- and ESX-1-dependent manner. *PLoS Pathog.*, **13**, e1006344.

77. Heidelberg,J.F., Elsen,J.A., Nelson,W.C., Clayton,R.A., Gwinn,M.L., Dodson,R.J., Haft,D.H., Hickey,E.K., Peterson,J.D., Umayam,L., *et al.* (2000) DNA sequence of both chromosomes of the cholera pathogen *Vibrio cholerae*. *Nature*, **406**, 477–483.

3 A REMARKABLY DIVERSE HORIZONTALLY MOVING GENE POOL IN ENVIRONMENTAL ISOLATES OF *VIBRIO CHOLERAE*

Drawing by Noémie Matthey, PhD

3.1 INTRODUCTION

The task of allocating prokaryotes into species has been historically problematic and based more in practicality rather than in a well-defined theory (1). Due to difficulties in defining ecological and genetic units of prokaryotes, theoretical interspecies boundaries are frequently blurry (1–4). This problem stems from the fact that prokaryotic genomes are often extremely fluid (2, 3, 5, 6). Comparative studies have demonstrated that most of genomic variation amongst strains from the same species does not derive from sequence variability, but rather from gene content (3, 7). In this context, the complete set of genes carried by all strains from the same species is referred to as the "pangenome", which can be subdivided into the "core genome" (genes present in all strains) and "accessory genome" (genes present in some but not all isolates) (7). The latter is often composed of genes found on mobile genetic elements (MGEs), which are blocks of genes associated to horizontal transmission (3, 8). In fact, a large mobilome, which corresponds to the totality of MGEs that can be transferred by horizontal gene transfer (HGT), is a typical feature of the prokaryotic genome landscape (5). MGEs such as by bacteriophages, genomic islands, plasmids, transposons and insertion sequence (IS) elements are highly complex and often mosaic in their composition (5, 9–11). They can change their chromosomal location within a cell or move from one organism's genome to another through HGT (transformation, conjugation or transduction) (5, 6, 12). The impact that HGT has on microbial evolution was made even more apparent with the increase in whole genome sequencing (WGS) projects (4, 6, 9, 13).

The importance of MGEs in microbial evolution was initially perceived due to their role in pathogenicity (3, 9, 11). Large genetic elements carrying virulence factors-encoding genes and found only in pathogenic isolates were initially referred to as "pathogenicity islands". The acquisition of such elements by bacteria allowed the shift from an environmental to a pathogenic lifestyle (3, 5, 11, 14, 15). Additional evidence over the years

clarified that these large islands are actually not solely restricted to pathogens and can also carry virulence unrelated genes. Therefore, these elements are now called by the broader term “genomic islands” (9). Genes carried in these islands can influence many bacterial adaptive traits besides pathogenicity, such as degradation of certain compounds, niche utilization, symbiosis, competition, defense against non-self-DNA, and many others (5, 9, 11). Genomic islands can be very complex in their composition, but usually have a skewed GC content compared to the rest of the genome. Moreover, they are often adjacent to tRNA genes or other strongly conserved genome regions and flanked by repeat sequences. Additionally, even though these islands are chromosomally inserted, they often times contain active integrases or transposases making them unstable and mobilizable (9).

Vibrio cholerae is an excellent model to study the evolutionary effects of MGEs. *V. cholerae* is a Gram-negative gamma-proteobacterium responsible for the acute diarrheal disease cholera, which has afflicted the world in seven pandemics since 1817. However, the primary habitat of these bacteria is the aquatic environment, where they often live attached to chitinous surfaces. Furthermore, most *V. cholerae* strains are not associated to cholera disease, and are referred in this manuscript as “environmental isolates”. In fact, only two (O1 and O139) out of the 200 *V. cholerae* serogroups are directly linked to cholera past and ongoing pandemics (16–19). Importantly, the acquisition of horizontally transmitted elements was pivotal in the evolution of the pandemic lineage (20). Among these elements is the CTX prophage, which carries the genes (*ctxAB*) encoding the main virulence factor, cholera toxin (CTX) (21). Additionally, these strains harbor two *Vibrio* pathogenicity islands (VPI): VPI-1 and VPI-2. VPI-1, also known as “TCP island”, is a 41-kb element that carries, among others, genes encoding the toxin-coregulated pilus (TCP), which is the second core virulence factor. TCP is not only essential for efficient intestinal colonization (22), but also serves as receptor for the CTX bacteriophage (CTXΦ) and therefore enables toxigenic

conversion (21–23). *Vibrio* pathogenicity island 2 (VPI-2) is a 57.3-kb island that encodes a type I restriction-modification (R-M) system and a Zorya defense system. It also carries the "*nan-nag* region", which encodes sialic acid transport and catabolism proteins as well as a neuraminidase (24). Neuraminidase is important for the unmasking of GM1 gangliosides in the intestine, which themselves are the receptors of cholera toxin (24–26).

The 6th and 7th (i.e., ongoing) cholera pandemics were provoked mainly by O1 serogroup strains from the classical and El Tor biotypes, respectively (19, 27). Besides genome sequence variability, these biotypes can be differentiated by variation in the mobilome landscape. For instance, they contain different CTX prophages (classical strains have CTX^{cla} and El Tor have CTX^{ET}) (27). Furthermore, El Tor strains have acquired two seventh pandemic islands (VSP), VSP-I and VSP-II (26, 28–30). VSP-I is a 16-kb region and most of its genes encode hypothetical proteins. However, it does carry a gene encoding a novel class of dinucleotide cyclases (DncV), which, in *V. cholerae,* is important for efficient host colonization and chemotaxis repression (30). More recently, *dncV* was shown to be part of a 4-gene phage defense system suggested to be the ancestor of the cGAS-STING antiviral pathway found in eukaryotes (31). VSP-II is a 27-kb region and has genes encoding mainly hypothetical proteins, but also proteins related to DNA repair and chemotaxis (25, 29). Contrary to VSP-I, there is no experimental evidence so far that qualifies VSP-II as a *bona fide* pathogenicity island (32). Apart from these common elements, the El Tor lineage has accumulated additional genome modifications over the years so that the strains can be classified into three waves of global propagation (27, 33). One key difference between waves 1 and 2/3 strains was the acquisition of a large MGE called SXT (an integrative and conjugative element), which mainly contains antibiotic resistance genes and phage defense features (33, 34).

There have been reports of *V. cholerae* strains causing localized cholera outbreaks that didn't reach pandemic levels (17, 35), as well as environmental TCP- and/or CTX-positive strains (36–40). These data demonstrate that the mere presence of cholera-causing MGEs is not deterministic for "pandemicity". Furthermore, while the mobilome of pandemic *V. cholerae* strains is well-characterized, the diversity hidden in environmental isolates remains poorly studied. The previously mentioned screenings (36–40) were PCR-based and conducted in a comparison- rather than discovery-driven manner. Few other studies have used whole-genome sequencing (WGS) to perform a deeper characterization of individual genomic islands in environmental *V. cholerae* strains and discovered important components, such as an auxiliary cluster of the type VI secretion system (T6SS) and a CRISPR-Cas module in another strain (41, 42).

We have recently performed long-read PacBio-based WGS of fifteen environmental *V. cholerae* isolates (43). These strains were originally collected in a non-endemic area in coastal California and classified into four clades (A-D) based on comparative genome hybridization (44, 45). As their characterization was made by microarray hybridization-based comparison against the genome of the pandemic strain N16961 (46), novel MGEs were not unveiled. In the present study we used the closed genomes of these strains to identify and characterize their genomic islands. We uncovered a remarkable mobilome diversity, including defense systems against foreign DNA, antibacterial modules, operons related to metabolic niche utilization, siderophore biosynthesis, osmoprotection, and also elements encoding proteins with similarity to TCP. Our investigation demonstrates the impressive variability encountered in the horizontal gene pool of *V. cholerae* and the potential that this could have for the continuous evolution of different lineages of this bacterium.

3.2 RESULTS, DISCUSSION AND EXPERIMENTAL OUTLOOK

3.2.1 Genomic islands in the environmental isolates

In order to investigate the genomic islands carried by environmental strains, we aligned each of their two chromosomes to the corresponding chromosome from a 7th pandemic representative El Tor strain, A1552 (47). Strains analyzed and their genome accession numbers can be found in Supplementary table 3.1. To start with, we inspected the strains for potential islands carried by the environmental isolates at the same genomic location as those islands specific to 7th pandemic strains: VPI-1, VPI-2, VSP-I and VSP-II. Importantly, we did not encounter any environmental isolate that carried pandemic strain-like islands *per se*, and none of the strains carried an island at the genomic location of VSP-II. Furthermore, we also examined the entirety of chromosomes 1 and 2 from these isolates and found additional genomic islands (Figure 3.1 and Supplementary table 3.2).

Many of the environmental genomic islands carry hypothetical genes or genes that are frequently found in other MGEs such as those encoding toxin-antitoxin (TA) modules, proteins involved in DNA repair and SOS response, and, most importantly, restriction-modification (R-M) systems. Some examples are the islands found is the same location of VPI-1 in strains SL6Y, SA3G and SA5Y (Figure 3.1). Notably, during the PacBio-based SMRT sequencing approach of these strains, information regarding putative DNA modifications was simultaneously recorded. With the help of Nicolas Guex and Christian Iseli of the Bioinformatics Competence Center facility of UNIL/EPFL (BiCC; Lausanne, Switzerland), we analyzed this data. As can be seen in Supplementary figure 3.1, many modified motifs were found in these strains. Several of these modified motifs were observed exclusively in one or few strains and are probably derived from methylating enzymes encoded on genomic islands. We also observed cases of marks that were restricted to certain locations of the genome, such as the integron island (data not shown). These

findings will be further studied in future projects related to genomic landscape and genome stability.

In the present *in silico* study, however, we focused on the characterization of islands with a more diverse genetic content. The islands that we characterized were classified into five main categories based on their assumed encoded functions: (I) antiphage defense systems; (II) bacterial antagonism weapons; (III) cell appendages (type IV pili); (IV) metabolism and (V) interactions with host. The locus tags that delimit all the characterized genomic islands can be found in Supplementary table 3.2. Furthermore, we selected six environmental isolates and assessed their transcriptomic profiles at low and high cell density (2 h and 6 h of growth) (Supplementary table 3.3). This allowed us to check whether the genes carried by these genomic islands were expressed and if that transcription or mRNA stability was quorum-sensing (QS) regulated. Finally, as this is an ongoing project, we describe the foreseen experiments for future experimental characterization of the selected genomic islands.

3.2.2 Antiphage defense systems

Bacteriophages are viruses that infect bacteria. The extent of the burden they inflict on bacteria has become more apparent by recent advances in computational methodologies and mining of public microbial genomes demonstrating their wide distribution (48). Nonetheless, bacteria have not remained passive against this predation pressure. These approaches have also exhibited the extensive variety of antiphage defense systems hidden in bacterial genomes, which have been reviewed in detail elsewhere (49–51). Even though we did not employ a computational prediction tool to identify defense systems in the environmental genomic islands, we did, through manual inspection, encounter a

considerable variety of them. Besides the broad distribution of R-M systems in all genomes, as previously mentioned, we also encountered strains with a Dnd defense system, a *qatABCD* cluster and CRISPR-Cas modules, as explained below.

3.2.2.1 Dnd defense system

Strains W10G (clade A), W6G, W7G and L6G (clade C) harbor a very similar island of 18 to 21 kb, inserted close to the same tmRNA (*ssrA*) where the VPI-1 is located in pandemic *V. cholerae* strains. These islands harbor a Dnd defense system (Figure 3.2a; locus tag numbers in Supplementary table 3.2), which was recently described as an antiphage immune system in bacteria and archaea. A DNA-incorporated phosphorothioate (PT) mark is used to discern self from foreign DNA, similar to R-M systems which are based on methylation. The system is usually composed of three parts: (i) The DndACDE proteins form a complex that acts as the modification element that carries out the PT modification in a sequence-specific manner; (ii) DndB is a transcriptional repressor that regulates the *dndACDE* cluster and therefore the PT level; (iii) the DndFGH proteins work as the restrictive element by recognizing and eradicating non-PT-modified and therefore non-self-DNA (52–55).

W6G and W7G harbor a completely identical island (100% by pairwise comparison), consistent with these strains being clonal (43). The respective island in strain W10G is 90% identical to the former two strains at the DNA level, with differences lying in the interspersed genes. The respective island in strain L6G is 67% and 64% identical to that carried by W6G/W7G and W10G, respectively. The variation is found mainly in the regulation and modification part of the cluster (*dndBCDE*), besides the absence of some Dnd-unrelated genes (marked with asterisks in Figure 3.2a). Furthermore, all four environmental strains

contain the complete set of *dnd* genes of the system, except L6G, which lacks *dndA*. This gene can be found separated from the *dndBCDE* operon in other bacteria or even absent from the genome (55), which is the case for L6G. Importantly, however, DndA is a cysteine desulphurase similar to the pyridoxal 5'-phosphate (PLP)-dependent cysteine desulphurase IscS. This protein is able to functionally replace DndA in the PT modification pathway when *dndA* is absent (56). L6G contains *iscS* (HPY16_09685 in chromosome 1) – as the other strains – and we therefore speculated that IscS might be working alongside DndBCDE. However, our DNA modifications analysis based on the SMRT sequencing contradicted that speculation (discussed below).

Expression profile

We evaluated the transcriptome of strains W10G and L6G when grown exponentially or to stationary phase. The Dnd defense system seems to be transcriptionally active in both strains (Supplementary table 3.4). Interestingly, especially in strain W10G, it seems that the modification cluster (*dndBCDE*) is more active at exponential phase, while the restriction cluster (*dndFGH*) is expressed at similar levels during both growth phases. This pattern would agree with the cell's needs, given that during exponential growth more DNA is synthesized, which needs to be modified.

Genomic DNA modification landscape

As mentioned previously, we obtained data regarding the DNA modification landscape of the environmental strains, which was analyzed by the BiCC facility. As shown in Supplementary figure 3.1, there was one motif (GGCC) that was found modified in its second base (with a not-identified modification) only in strains W10G, W6G and W7G. While the Dnd system does not directly modify the DNA bases, the modification of the phosphate backbone could impact the SMRT sequencing velocity and therefore change the inter-pause sequencing time, which is then marked as a DNA modification of unknown function. It is

therefore tempting to speculate that this mark refers to the Dnd defense system-derived PT modification. Its absence from strain L6G might be explained by the lack of *dndA* in the cluster. In strains W10G, W7G and W6G, the modification was present in around 26-27% of the overall motif, and the marks were found well distributed throughout the 2 chromosomes and the megaplasmid that is found in strains W7G and W6G (Supplementary figure 3.2 and data not shown).

Experimental outlook

Plasmid electroporation seems less efficient in strain W10G, hinting to a restriction mechanism. We will therefore test the impact of deleting the restriction cluster in plasmid electroporation efficiency in this strain. As for the modification cluster, we will genetically engineer the WT strain to encode a FLAG-tagged version of DndB and then check its protein production by Western blot analysis. Moreover, we will transfer the *dndBCDE* cluster into our 7th pandemic *V. cholerae* type strain, A1552, and *E. coli,* followed by the isolation of plasmids from these strains. These – in theory – PT-marked plasmids will then be electroporated into W10G to check whether they are protected from degradation by the restriction cluster. Finally, we will use SMRT sequencing to check whether a W10G variant lacking the Dnd defense system has lost the DNA modification described above (Supplementary figure 3.2).

3.2.2.2 CRISPR-Cas

Strains SP7G, SA10G, and SL4G harbor CRISPR-Cas modules on different genomic islands (Figure 3.2b). CRISPR (Clustered Regularly Interspaced Short Palindromic Repeats) is a type of immune system that is widespread in bacteria and archaea (57, 58). The “immune memory” is collected in CRISPR arrays where short DNA sequences (spacers)

derived from previously infecting pathogens are kept separated by conserved repeats (58, 59). The array is accompanied by a regulatory leader sequence and genes encoding Cas (CRISPR associated) proteins, which are responsible for the immunity function (58). Briefly, the mechanism of CRISPR immunity includes three stages. The adaptation phase corresponds to the acquisition of new spacer sequences derived from targets (such as bacteriophages) and their insertion into the CRISPR array. In the expression phase, the CRISPR array is read and processed into small RNAs (crRNAs), which are complexed by Cas proteins. Finally, the interference stage consists of the crRNA-Cas complex searching the cell for potential foreign nucleic acids with complementarity to the crRNA, which induces target cleavage (57–60). Arms-race evolution has led to the diversification of Cas proteins and the CRISPR systems as a whole, which were classified into subtypes (extensive details on their classification can be found, for instance, in 67, 69, 70).

Strain SL4G harbors a class 1 subtype I-C CRISPR-Cas module (59) on a 92-kb island inserted in chromosome 2. The system contains genes that are annotated as *cas3*/*5*/*8*/*7*/*4*/*1/2*, followed by a repeat/spacer array with 28 predicted repeats (Figure 3.2b; right panel; locus tags in Supplementary table 3.2). On the other hand, strains SP7G and SA10G harbor a very similar island inserted next to *ffs* ncRNA (signal recognition particle sRNA) on chromosome 1. While in SP7G the island has a size of 37.5 kb, the one of SA10G it solely 25 kb in length. Most parts of the island (roughly 20 kb) are almost identical and include the CRISPR-Cas module (Figure 3.2b; left panel; locus tags in Supplementary table 3.2). This module fits with the previously described minimal version of a subtype I-F (59, 63, 64), which was observed in several *Vibrio* species (65) and consists of a transposase gene (*tniQ*), *csy2 (cas5)*, *csy3 (cas7)*, *cas6* and a repeat/ spacer array, which contains 4 and 3 repeats in SP7G and SA10G, respectively (Figure 3.2b; left panel).

Expression profile

Strains SP7G and SA10G were checked for their transcriptome at exponential and stationary phase. In both strains their minimal subtype I-F CRISPR-Cas modules seem to be transcriptionally active, especially in stationary phase (Supplementary table 3.5).

3.2.2.3 QatABCD defense system

Strains SL5Y and W10G both contain a QatABCD defense system. The *qatABCD* operon was initially discovered by Gao *et al.*, when the authors predicted putative defense systems in bacterial and archaeal genomes that were available in GenBank (by the end of 2018), independent of domain annotations (66). The operon was identified in the genome of an *E. coli* strain and, after cloning and transforming it into a surrogate *E. coli* lab strain, it was shown to be effective as a defense against phages P1, λ and T3 (66). The operon consists of an ATPase (named QatA), a hypothetical protein QatB, the 7-cyano-7-deazaguanine synthase QueC (renamed QatC in the cluster) and a TatD nuclease (renamed QatD). The ATPase QatA and the nuclease QatD appear to be the most relevant components of the system, as mutations in their active sites significantly reduced their antiphage defense against λ (1.5-log for QatA and 3-log for QatD). Deletion of *qatB* or the mutation of active site residues of QatC had no effect in phage defense efficiency (66).

In strains SL5Y and W10G, the *qatABCD* genes are located on different genomic islands when the two strains are compared. In strain SL5Y, the ~5-kb operon is located on a 15-kb island inserted in the same location where the VSP-I is located in pandemic *V. cholerae* strains, which is close to the *csrB RNA* gene. Besides *qatABCD*, the island also carries genes encoding an integrase, an ABC-transporter, and several hypothetical proteins (Figure 3.2c; locus tags in Supplementary table 3.2). In strain W10G, the operon is located

on a 26.6-kb island inserted next to a Asn-tRNA on chromosome 1, which also carries genes encoding a contact-dependent inhibition (CDI) system (see below) (Figure 3.2d; locus tags in Supplementary table 3.2). In a pairwise nucleotide comparison, the *qatABCD* clusters carried by SL5Y and W10G turned out as 59% identical. Pairwise comparisons of the translated proteins from each strain gave identities between 46.8% (QatB) and 58.6% (QatC).

Expression profile

The *qatABCD* defense system seems to be transcriptionally active in both strains (Supplementary table 3.6). While the expression level of the operon is 2 to 4-fold higher in stationary phase in strain SL5Y, this expression increase is less pronounced in strain W10G.

Experimental outlook

We will clone the operon under the control of an inducible promoter in *E. coli* and test its role in antiphage protection.

3.2.3 Bacterial antagonism weapons

Besides dealing with constant pressure from bacteriophage infections, bacteria also have to compete with one another, especially considering that they are usually embedded into biofilms and complex communities where the resources can be scarce (67). In this sense, research advances have shown the importance of antagonism strategies for bacterial survival and the variety of mechanisms that can be employed by one single strain (68, 69).

In the case of *V. cholerae*, the focus on interbacterial competition strategies relies on the type VI secretion system (T6SS) (68), a multiprotein complex that assembles across the double membrane of Gram-negative bacteria and can puncture target cells, delivering toxic molecules (70–72). All *V. cholerae* strains have the gene clusters encoding the T6SS (73).

However, specific signals are necessary for the machinery's induction in pandemic strains (74, 75), while non-pandemic *V. cholerae* keep the T6SS constitutively active (40, 43, 76, 77). We recently demonstrated that the environmental isolates investigated in this study also have a constitutively active T6SS and employ an astonishingly diverse effector repertoire to fight bacterial competitors and eukaryotic grazers (43). We also discovered that clade C strain SP7G contains an additional auxiliary T6SS cluster (Aux 6), which is located in a genomic island on chromosome 2 and is active under common laboratory conditions (43) (Figure 3.1). In the current characterization of genomic islands carried by these environmental isolates, we discovered gene clusters encoding two other antibacterial systems: a contact-dependent inhibition (CDI) system in strain W10G and a S2 pyocin in strain SA10G, which will be discussed below.

3.2.3.1 Contact-dependent inhibition (CDI)

In the same island in strain W10G on which the *qatABCD* antiphage defense system gene cluster is located, a contact-dependent inhibition (CDI) system is also encoded (Figure 3.2d; locus tags in Supplementary table 3.2). The CDI cluster found in strain W10G (*cdiBAI*) follows the same architecture as is found in most α-, β-, and γ-proteobacteria (78). *cdiB* and *cdiA* encode a two-partner secretion system, whereby CdiB is a protein that fosters the export of CdiA. CdiA is a very large protein that is exported through the CdiB pore as a filament that extends out of the cell, while the C-terminal toxic domain (CdiA-CT) remains inside the bacterial periplasm (79). Binding of CdiA to a receptor in the target bacterial cell (as, for instance, the ubiquitous proteins BamA or OmpC/F) fosters the complete secretion of the protein with the C-terminal domain being translocated into the target cell's periplasm (78, 79). *cdiI* encodes the cognate immunity protein that protects siblings (80). The C-terminal domains of CdiA are highly diverse and include metallopeptidases, nucleases,

ADP-ribosyl cyclases, and adenosine deaminases, to name a few (78, 80, 81). Each CdiI immunity protein is associated to a specific CdiA toxin (CdiA-CT) (81), and these encoding gene pairs can be horizontally transmitted (80). Interestingly, recent studies have shown that the CDI system also serves community and cooperative behaviors such as biofilm formation, in a phenomenon termed contact-dependent signaling (78, 82, 83).

Expression profile

CDI was initially discovered in an *E. coli* strain that kept the system constitutively active (84), although many species tightly regulate their CDI production (80, 85). Our RNAseq data from strain W10G indicates that its CDI system is active under laboratory conditions with a trend (~2 to 5-fold) for higher expression at exponential phase (Supplementary table 3.7).

Experimental outlook

To verify that the transcriptomic data adequately correlate with the protein levels, we will FLAG-tag CdiB and monitor its production by Western blot analysis. Furthermore, we will construct mutants and test the system's functionality in competition assays (86) against the WT W10G strain, other *V. cholerae*, other Vibrios, or *E. coli*. If the system turns out to be indeed active, we will transfer the immunity gene with an inducible promoter into other *V. cholerae* strains and test for strain's protection.

3.2.3.2 S2 pyocin

Strain SA10G carries a 19.2-kb long island on chromosome 1, inserted next to a predicted cobalamin riboswitch. The island is composed mainly of hypothetical and phage genes, but also contains a gene annotated to encode a S2 pyocin (*pys2*) and its corresponding immunity gene (a function that BLAST analysis confirmed; Figure 3.2e). Pyocins are

bacteriocins from *Pseudomonas aeruginosa* which are active against closely related strains due to the presence of specific receptors required for their toxicity (87, 88). Three types (R-, F-, and S-type) of pyocins have been identified in *Pseudomonas*. Types R- (rigid) and F- (flexible), also referred to as tailocins, resemble bacteriophage tails and are resistant to nucleases and proteases. R- and F-type pyocins are synthesized by complex gene clusters that include lysis cassettes, which allow their release from the producer by cell lysis (88, 89). On the other hand, S-type pyocins are soluble protease- and heat-sensitive particles that show no homology to bacteriophage tail proteins (88, 90). Instead, S-pyocins contain a toxic domain, which is neutralized by a protective immunity protein during secretion. The complex is then released to allow toxic activity upon binding of the toxin to specific receptors on target cells (88, 90).

The DNAse toxic activity of the S2 pyocin from *P. aeruginosa* is conferred by a C-terminal HNH domain, which is also found in the predicted protein encoded by the annotated S2 pyocin gene in strain SA10G (88, 90, 91). Additionally, the S2 pyocin of strain SA10G harbors a translocation domain pyocin S (pfam06958) and a LysM motif, involved in peptidoglycan binding (92). This domain is not found in the S2 pyocin from *P. aeruginosa* PAO3, which contains a SMC (structural maintenance of chromosomes) domain instead. Notably, closer inspection of the S2 pyocin protein sequence by BLASTp suggested that it might belong to the class of MIX (marker for type VI effectors) effectors, as a MIX-V domain was predicted in its N-terminus. Importantly, *V. cholerae*'s T6SS is able to deliver MIX effectors such as the VasX protein of pandemic strains (93). For this reason, we speculate an alternative secretion mode for this gene product compared to *bona fide* S pyocins, namely via its interaction with the T6SS machinery.

Expression profile

In *P. aeruginosa*, expression of all three pyocin types (R-, F- and S-) is regulated by the PrtN activator protein, which binds a regulatory sequence (P-box) located upstream of the pyocin cluster (88). Furthermore, pyocin production seems to be induced under SOS response-triggering stress conditions, such as UV radiation or mitomycin C treatment (87, 88, 90). Our RNAseq data indicates that the genes encoding the S2 pyocin and its immunity protein are expressed in strain SA10G, with slightly higher expression at high cell density compared to low cell density (Supplementary table 3.8). It will be interesting to test experimentally whether the system is indeed active.

Experimental outlook

Similar to what is planned for the CDI, we will FLAG-tag the S2 pyocin and check the protein production by Western blot analysis. Next, a mutant lacking the pyocin and immunity genes will be constructed to test the pyocin's killing activity. Precisely, survival of this mutant will be scored when exposed to conditioned medium derived from the wild-type strain and also from a T6SS knockout strain, to inspect whether the toxin is T6SS-delivered. Furthermore, we will also engineer an inducible version of the cluster in our lab strain (A1552) and test its subsequent competitiveness against its parent and other *V. cholerae*, other *Vibrio* species, or non-*Vibrio* bacterial strains.

3.2.4 Cell appendages: type IV pili

Bacteria are most commonly found organized as biofilms attached to surfaces (67). In order to sense and appropriately respond to these surfaces and environmental conditions, they take advantage of complex cell appendages called type IV pili (T4P) (94). These nanomachines are composed of a conserved set of proteins that assemble surface-exposed filaments formed primarily by major pilins (95). T4P can be involved in a variety of functions,

such as attachment to biotic and abiotic surfaces, biofilm formation, twitching motility and DNA-uptake (94, 95). Importantly, many of these cell appendages are also key virulence factors in pathogenic bacteria (96).

V. cholerae can be found in diverse environments, ranging from the human small intestine to chitinous surfaces in the aquatic environment (97, 98). The diverse nature of *V. cholerae*'s lifestyles requires specific adaptations to interact with these surfaces, which is reflected in a diverse set of T4P. Pandemic *V. cholerae* commonly encode three T4P systems. First, the DNA-uptake pilus is a T4aP (A-type prepilin) produced when *V. cholerae* grows on chitinous surfaces, allowing DNA-uptake and natural transformation (99–101). More recently, this pilus was also shown to be important for chitin colonization and kin recognition (102). Second, mannose-sensitive haemagglutinin (MSHA) T4aP are important for sensing and attachment to surfaces (103–106). Finally, the toxin co-regulated pilus (TCP) is a T4bP (B-type prepilin). As mentioned previously, TCP is a fundamental virulence factor of toxigenic *V. cholerae*, as it is crucial for auto-aggregation and intestinal colonization but also serves as the receptor for the CTXΦ (21–23).

Our analysis of the environmental strains' genomes demonstrated that a subset of strains contain T4P loci in genomic islands. As explained in detail below, three strains contain a large genomic island that is located at the same location as VPI-2 in pandemic *V. cholerae* strains and encodes an operon for a TCP-like pilus. Three other strains contain an island with the entire operon encoding the production of a Tad pilus, which had not been observed in *V. cholerae* so far.

3.2.4.1 Toxin-coregulated pilus (TCP)

Strains W10G (clade A), W6G, and W7G (clade C) carry a very large (101-105-kb) island at the same location as VPI-2 in pandemic strain A1552, next to a tRNA-Ser gene. The island contains several hypothetical genes, but also homologues of genes related to DNA repair and SOS response, toxin-antitoxin systems, metalloproteases, efflux pumps, etc (Figure 3.3a; locus tags in Supplementary table 3.2). The island also contains many *tra* genes that are involved in the production of the type IV secretion system-related F pili, which are involved in conjugation. F pili allow bacteria to establish contact and form a stable mating pair. Pilus retraction then permits that cells come close together, when the DNA transport into the recipient cell can commence through a transfer pore (5, 107–109). The presence of these genes in the island indicates its ability to move horizontally by conjugation (e.g., potentially an integrative and conjugative element).

Interestingly, this island also encloses a cluster of *tcp*-related genes that most likely code for the production of a TCP-like pilus. PCR-based genetic screens of *V. cholerae* samples collected around the world have previously exposed environmental strains carrying the major pilin gene *tcpA* and/or the regulatory gene *toxT*, but the presence of the rest of the operon was never determined (36–38). The gene organization of the environmental *tcp* operon strongly resembles the organization of the pandemic *tcp* cluster, even though they only share 60% pairwise nucleotide identity. The clusters carried by W6G/ W7G (clonal strains) and W10G are 95% identical in a pairwise comparison. Most of the divergence between their operons lies in the major pilin gene *tcpA* and the initial region of the *tcpB* gene. In pandemic *V. cholerae*, the *tcp* genes are activated by the transcriptional regulator ToxT/TcpN (110), a homolog of which is also encoded within the environmental *tcp*-like operons. Protein alignment of the pandemic ToxT with W10G's ToxT homolog shows only 27% identity; however, the predicted protein contains the typical AraC domain of

transcriptional regulators, alike the pandemic ToxT protein. Notably, *toxT* is activated by ToxR/S and TcpP/H in pandemic strains, which themselves respond to environmental signals such as bile and osmolarity, for example (110–114). While *toxR* and *toxS* are chromosomally encoded and conserved in the environmental isolates, *tcpP* and *tcpH* are part of the VPI-1 island in pandemic strains and absent in the environmental genomic island.

Expression profile

Our RNAseq data suggests that the *tcp*-like genes are poorly expressed in strain W10G (Supplementary table 3.9) with slightly higher transcript levels at low cell density. These low transcript levels suggest that environmental stimuli might be required to activate the gene cluster, most likely through the ToxT-like regulator.

Experimental outlook

As the *tcp* operon is apparently silent in strain W10G, we will first genetically engineer a strain with an inducible copy of the transcriptional regulator gene (*toxT*-like*)* to test the activation of the gene cluster. We will test the exchangeability of the ToxT-like protein against its pandemic counterpart (ToxT*)* to observe how conserved the regulatory circuits are. Indeed, we hypothesize that the DNA-binding and gene expression activation part of the regulator might be conserved for their function. Induction of the gene clusters will then be tested by qRT-PCR and Western blotting, after adding a FLAG-tag to the TcpA/TcpA-like pilins. Finally, we also aim at visualizing the environmental TCP-like pilus and its dynamics by microscopy, which will be done through a commonly used cysteine knock-in coupled to cysteine-labelling approach (115), which is frequently used in our group (101, 115, 116; and unpublished work from J. Pereira on *V. cholerae*'s TCP).

3.2.4.2 Tad pilus

Clade D strains E7G, SA7G, and SA10G have an identical 16-kb island inserted close to a cluster of 5 tRNA genes, 8 kb downstream of the VSP-I location in A1552. The cluster is mainly composed of genes encoding homologues to components of a tight adherence pilus (Tad, also called Flp or Fap) (Figure 3.3b; locus tags in Supplementary table 3.2). This is a T4cP derived from archaeal T4P (118) which is important in the lifestyle of several bacterial genera including *Pseudomonas*, *Haemophilus*, *Yersinia* and *Caulobacter* (119, 120). Moreover, its wide distribution throughout the bacteria suggests HGT-driven dissemination (121). Notably, the pilus of *C. crescentus* was the first Tad pilus to have its extension and retraction cycles visualized by microscopy using cysteine labelling (115). Furthermore, it was recently demonstrated that *tad* loci are widespread in the Vibrionaceae family, even though there was no evidence of *tad* genes being present in *V. cholerae* strains (122). *V. vulnificus*, for instance, contains three *tad* clusters. The third cluster (*tad-3* or *iam*) was shown to be involved in aggregation, biofilm formation under hydrodynamic flow, and oyster colonization (122, 123).

The *tad* cluster found in clade D strains seems to contain genes encoding all proteins required for Tad pilus assembly and function. We found homologues encoding the secretin RcpA, both TadA and TadZ ATPases, the inner membrane platform proteins TadB/C, the major pilin Flp and minor pilins TadE/F, assembly proteins TadG and RcpB, prepilin peptidase TadV (which is absent from *tad-3* locus from *V. vulnificus*) and the pilotin TadD (118). Manual inspection of the cluster organization indicates that it resembles the *tad-1* cluster from *V. vulnificus* (122).

Expression profile

Our RNAseq data indicates that most genes from the *tad* cluster found in clade D strains are expressed at low levels, at least in strain SA10G (Supplementary table 3.10). Genes

encoding the assembly protein TadG and the prepilin peptidase TadV were more highly expressed than the other genes, and that expression was around 2-fold higher at high cell density compared to low cell density.

Experimental outlook

In order to check the protein expression of the Tad pilus, we will FLAG-tag the inner membrane platform protein TadB and check its production by Western blotting. Moreover, we previously observed that strain SA10G grows on certain solid surfaces (such as SM/5 media used in *Dictyostelium discoideum* amoebae plaque assays; see chapter 4) with a very distinctive and rather opaque aspect compared to all the other environmental isolates that don't encode a Tad pilus; we therefore wondered whether this correlation could reflect a Tad pilus-related phenotype. To check this idea, we will interrupt the *tadB* and/or *flp* gene(s) in strain SA10G and compare the mutants' phenotype to that of the WT upon growth on those surfaces. Furthermore, we will image the Tad pilus production and dynamics by microscopy using the above-mentioned cysteine knock-in and labelling approach (102, 115–117).

3.2.5 Metabolism

HGT has an immense impact on bacterial evolution due to the mobilization of operons encoding physiological traits that allow exploration of new environments, leading to ecological innovation (5, 9, 12). Notably, we observed a striking diversity of islands with metabolic-related traits in the environmental strains, which could provide a rich source of investigation. For the sake of briefness, we will focus on two such islands: an operon that we suggest encodes for N-Acetylgalactosamine utilization and a widespread island containing an operon for siderophore biosynthesis and osmolyte production.

3.2.5.1 N-Acetylgalactosamine-linked PTS system

Clade B strains SA5Y, SL4G, and SL5Y harbor a 9-kb island (98% identical to each other) that is located close to a putative signal recognition particle RNA (*ffs* ncRNA) on chromosome 1. Prokaryotic Genome Annotation Pipeline (PGAP) automatic annotation of these genes was confusing regarding the sugar specificity, which is often the case. However, upon closer inspection via BlastP and comparison with clusters in the literature, we identified the cluster as one encoding a putative N-Acetylgalactosamine (Aga) amino sugar utilization system (Figure 3.4a, locus tags in Supplementary table 3.2). The enzyme II (EII) for Aga utilization belongs to the mannose and L-sorbose family as it contains, besides the usual subunits IIA, IIB and IIC, a fourth, IID, subunit (124). Subunits IIB, IIC and IID are encoded by genes *agaV*, *agaW* and *agaE*, respectively. The IIA domain, encoded by *agaF*, is shared by EII for Aga and D-galactosamine (Gam). The gene organization found in the here-described strains deviates slightly from the well-studied cluster in *E. coli*, which contains an operon for both Aga and Gam utilization (125). Instead, the cluster found in environmental *V. cholerae* matches perfectly the organization found in *V. vulnificus* (strain CMCP6) (126). The main difference of these *Vibrio* clusters is that they contain two transcriptional repressors AgaR instead of solely one in *E. coli*.

Notably, at the time of their collection in California, all environmental isolates were tested by Keymer and colleagues for their growth in different carbon sources using phenotypic microarrays (Biolog, Hayward, USA). SA5Y, SL4G, and SL5Y were able to grow on N-Acetylgalactosamine, contrary to all other tested strains of this collection (44). For this reason, we concluded that the clusters encode N-Acetylgalactosamine utilization systems.

Expression profile

Our transcriptomic data shows that the PTS-Aga system from strains SA5Y and SL5Y is poorly expressed (Supplementary table 3.11). This is somewhat expected, as these strains

were grown in rich LB medium without supplementation of the inducing amino sugar and the likely occurrence of catabolite repression. Consistent with the lack of expression, one of the repressor genes was highly expressed in both strains SA5Y and SL5Y, suggesting that AgaR actively represses the cluster.

Experimental outlook

As mentioned previously, there is data demonstrating that solely these three strains are able to utilize N-Acetylgalactosamine (44). We will therefore delete the EIIBCD subunits in SA5Y and test whether the resulting strain can no longer grow in minimal medium supplemented with N-Acetylgalactosamine as sole carbon source. We will also check whether it is possible to transfer the PTS cluster into our lab strain A1552 by natural transformation and whether such a transfer confers the bacteria with the ability to catabolize N-Acetylgalactosamine.

3.2.5.2 Piscibactin biosynthesis and osmoprotection cluster

All strains from clade C and D harbor the same 39-kb island inserted on chromosome 2, making this the most widespread genomic island that we have identified in the here-described collection of 15 environmental *V. cholerae* strains (Figure 3.4b, locus tags in Supplementary table 3.2). The presence of genes encoding putative iron metabolism-related proteins such a TonB-dependent receptor and two large (~6-kb and 12-kb) nonribosomal peptide synthetase (NRPS) in the island indicated that the cluster could be related to siderophore biosynthesis. Siderophores are iron-scavenging molecules that are produced and secreted by bacteria, fungi and plants. After chelating available iron in the environment or inside a host, these siderophores re-enter the producing cell via specific receptors and are metabolized to release the needed iron inside the cell (127–129). After visual inspection of a recent compilation of siderophore clusters in Vibrionaceae (128), we observed that the

cluster found in the here-described strains resembles a widely distributed piscibactin biosynthetic cluster. This cluster was first recognized and studied in a highly virulent *Photobacterium damselae* subsp. *piscicida* strain DI21. Notably, the biosynthetic cluster was shown to resemble the yersiniabactin-encoding gene cluster that is located on the *Yersinia* high-pathogenicity island (HPI) (130). In *P. damselae* subsp. *piscicida* DI21, the piscibactin cluster is located on a transmissible plasmid (pPHDP70). It was demonstrated that this plasmid could be conjugated into diverse gammaproteobacterial species through the help of the conjugation machinery encoded on the co-resident plasmid pPHDP60 (130, 131).

We confirmed that the cluster found in the environmental *V. cholerae* isolates indeed corresponds to the piscibactin biosynthesis cluster by aligning the representative cluster from strain L6G to the pPHDP70 plasmid (accession number KP100338.1) (130, 131). Even though we observed high sequence variability (the clusters were roughly 50% identical at the nucleotide level and between 37-62% identical at the protein level), the gene content and order were identical. We therefore named the genes according to the nomenclature used in *P. damselae* subsp. *piscicida* DI21 (*irp* genes, for iron-regulated proteins). Furthermore, there are two genes encoding ABC transporters at the very end of the environmental genomic island, separated from the rest of the piscibactin biosynthesis cluster by a 5.7-kb glycine-betaine biosynthesis operon. These two ABC transporters, together with the TonB-like receptor (encoded by *frpA*), were proposed to be responsible for piscibactin transport (131).

In order to grow and survive under high osmolarity conditions, bacteria have evolved a coping mechanism that is based in the accumulation of osmolytes (compatible solutes) in their cytoplasm to keep the cell's turgor pressure constant (132, 133). Glycine betaine, carnitine and proline are the main osmolyte compounds used in bacteria for osmoadaptation (132). Though *de novo* glycine betaine synthesis is rare, bacteria can transport choline into

the cell and transform it into glycine betaine. Choline is transported into the cell using high affinity (BetT; from the Betaine/ Choline/ Carnitine-Transporter family - BCCT) or low affinity (ProU) transporters. Next, choline dehydrogenase (BetA) oxidizes choline to glycine betaine aldehyde. The glycine betaine aldehyde dehydrogenase (BetB) then converts it to glycine betaine. The *betTBA* operon co-localizes with *betI* in *E. coli* , which encodes a choline-sensing repressor protein (132). The same gene organization is found in the genomic island of the environmental isolates. Interestingly, pandemic strains of *V. cholerae* have a BCCT-family glycine betaine transporter called OpuD (also known as BetS; VC1279) (134), a homolog of which is also encoded in the environmental strains. However, pandemic strains do not harbor the *bet* genes for glycine betaine synthesis from choline. We therefore suggest that strains harboring the genomic island-contained *betTIBA* operon might have a fitness benefit when they encounter high osmolarity environments. Interestingly, it was previously shown that *V. parahaemolyticus* strain RIMD2210633 contains two compatible solute synthesis systems (*ectABC* - for ectoine synthesis - and *betABI*; each operon clustered with a gene encoding a ProU transporter), and also four homologues of BCCT transporters. This is much more than what is usually found in other *Vibrio* species. This extensive array was reflected in a growth advantage under different salinities and temperatures when compared to *V. vulnificus* YJ016, *V. cholerae* pandemic strain N16961, and *Aliivibrio fischeri* ES114 (135).

Expression profile

As this island is widespread in the environmental strains collection, we had RNAseq data from three strains: L6G, SP7G, and SA10G. All strains seemed to express the genes on this island, with a very clear difference in growth phase dependency for the different parts of the island. Precisely, the genes related to piscibactin biosynthesis were higher expressed at low cell density, with some genes showing 20-30-fold higher transcript levels at 2 h of growth

when compared to the samples after 6h of growth (Supplementary table 3.12). QS-dependent regulation of siderophore biosynthesis has been commonly observed, as it helps preventing the accumulation of siderophores above the cellular uptake capacity (136, 137). On the other hand, the *betTIBA* cluster for choline transport and conversion to glycine betaine was higher expressed at high cell density than at low cell density, which was the case for all three strains (Supplementary table 3.12). This finding is also in agreement with previous work in other species such as *V. harveyi* (now called *campbellii*), in which the QS-master regulator LuxR activates the expression of the *betIBA-proXWV* operon at high cell density. Interestingly, these authors also observed that the auto-repressor BetI was able to repress QS in *V. harveyi*, demonstrating a linked regulation of osmotic stress and QS (138).

3.2.6 Interactions with a host

Considering that genomic (pathogenicity) islands were initially identified for carrying virulence factors (3, 9, 11), the perhaps most expected components of these islands would be genes involved in interactions with a host. Indeed, we found examples of gene clusters with such characteristics. A large group of strains carried a *lux* operon, giving bioluminescence ability to the bacterium. Moreover, a large proportion of environmental isolates carried a VPI-2 variant island that encodes a type III secretion system (T3SS) and a *nan-nag* region. These systems will be further explained below.

1.2.6.1 *lux* operon

All clade D strains and also clade C strains SL6Y, L6G and W6G/ W7G harbor a 11 to 19-kb island on chromosome 1 that contains a *lux* operon, whose gene products are known to be involved in bioluminescence. The island also harbors a series of conserved genes with

diverse functions (related to histidine synthesis and nucleotide metabolism, among others) and genes that vary among strains (marked with asterisk in Figure 3.5a, locus tags in Supplementary table 3.2). Also, the genomic island of strain SA10G is slightly longer than the one of the other strains, as it contains three extra ORFs at the end of the *lux* gene cluster.

Bioluminescence is the chemical process of light emission, in which the enzymatic reaction is catalyzed by luciferase. Bioluminescence is particularly important in marine environments, where species of fish, crustaceans, dinoflagellates and bacteria are able to emit light. In these darker environments, bioluminescence can be useful for food location, mate attraction, prey allurement and predation defense (139, 140). Even though some animals are able to produce light themselves, other species rely on bacterial symbionts to outsource light production. Bioluminescent bacteria are usually found in the *Aliivibrio*, *Vibrio* and *Photobacterium* genera belonging to the Vibrionaceae family (141). All bacterial species rely on the same proteins for light emission, encoded by the *luxCDABE*(*G*) core operon (gene order might vary slightly, and some cluster might contain additional genes) (139, 140, 142). While the *luxAB* genes encode the luciferase subunits, *luxCD* and *luxE* encode a fatty acid reductase complex that is necessary for the substrate (long-chain aldehyde) synthesis. *luxG* encodes a flavin reductase, which provides the reduced flavin used by the luciferase enzyme for bioluminescence production (140, 142). Furthermore, bioluminescence production is frequently under quorum-sensing (QS) control and therefore best studied in *A. fischeri* (previously named *V. fischeri)* and *V. campbellii* (previously named *V. harveyi*) (142).

Previous analysis of intraspecies transfer of *lux* operons in Vibrionaceae demonstrated that the cluster is primarily vertically transmitted (141). This notion is supported by our work since strains that harbor the *lux* cluster are closely related (Figure 3.1). Furthermore, the discovery of bioluminescent *V. cholerae* species is not improbable. Indeed, Grim and collaborators (143) already tested bioluminescence production and

presence of the *luxA* gene in hundreds of *V. cholerae* strains collected in the Chesapeake Bay (USA) as well as in Bangladesh. They observed a high incidence of bioluminescence in non-O1/non-139 strains in the US, contrary to a lower incidence in Bangladesh. Furthermore, they found no bioluminescent or *luxA*-positive O1/O139 strains.

Expression profile

The transcriptomic data of strains L6G (clade C) and SA10G (clade D) show that the *lux* operon is higher expressed at high cell density, as expected (142), and that the genes encoding the luciferase subunits LuxAB were the most strongly expressed genes in the operon of both strains (Supplementary table 3.13).

Experimental outlook

Future foreseen experiments aiming at verifying the functionality of the gene products encoded on this genomic island will be, of course, bioluminescence measurements after growth in marine medium as previously reported (143). Furthermore, we will test the effect of deleting the cluster and also the QS regulator *hapR* under those conditions. We will also transfer the operon into our lab strain A1552 by natural transformation, as measuring luciferase could be a useful tool in the laboratory when engineered, for instance, as a transcriptional reporter construct.

1.2.6.2 *nan-nag* region and type III secretion system (T3SS)

All strains from clade B and also strains SL6Y and SP6G from clade C harbor an almost identical island next to a tRNA-Ser gene on chromosome 1, which is the same location occupied by VPI-2 in pandemic *V. cholerae* (Figure 3.5b; locus tags in Supplementary table 3.2). This island is around 65-67-kb long, out of which 13-kb is homolog to the pandemic *nan-nag* region, which encodes a sialic acid utilization pathway (24–26). The island of strain

SP6G (clade C) is longer (71-kb) than the one from the other environmental strains, as it also contains a 4-kb region with genes homologues to ORFs located on VPI-2 (VC1804 - VC1810 locus tags in pandemic N16961 strain (46)). Notably, this SP6G island lacks the putative prophage that is located between this 4kb region and the *nan-nag* region in pandemic VPI-2 (24).

The remaining part of these environmental islands contains several hypothetical genes but also includes a *tdh* gene, which encodes a homolog to the thermostable direct hemolysin of *V. parahaemolyticus* (Figure 3.5b). Indeed, Tdh has been well characterized in other *Vibrio* species and represents a major virulence factor in *V. parahaemolyticus*, due to its ability to cause apoptotic cell death (144, 145). Interestingly, in *V. parahaemolyticus* strains, *tdh* is often associated with the concomitant presence of a type III secretion system (T3SS) (146). Consistent with this co-occurrence, further analysis of these environmental islands identified several genes encoding homologs of T3SS proteins.

The NF-T3SSs (non-flagellar T3SS), also called “injectisomes”, share several features with flagella, and are thought to have evolved from them (147). The complex machinery is formed by more than 20 kinds of proteins and is assembled by different substructures. These include (i) cytoplasmic components of the sorting platform (ATPase complex and the cytoplasmic, C-ring); (ii) an inner membrane export apparatus; and (iii) a needle complex (flagellar basal body) that spans both inner and outer membranes enclosing an inner rod, a pilus-like appendage (needle) that spans between the bacterial and target host cells, and the translocon (needle tip) that is translocated into the host cell membrane. The T3SS assembly establishes a channel between the bacterial and host cells, allowing the transport of effector proteins that can modulate the host response. Although the structural components of the system are conserved, effectors display a high diversity in

function and structure, reflecting the variety of interactions established by T3SS-harboring bacteria and their hosts (some reviews include 80, 128–132).

The first identification of a T3SS in *V. cholerae* came about when a O39 serogroup strain, AM-19226, was isolated from a patient with severe diarrhea in Bangladesh. Although the strain lacked cholera toxin and TCP, it displayed considerably high pathogenicity against animals in laboratory studies (153). Sequencing of the strain allowed the identification of a cluster with high similarity to the T3SS-2 of pandemic *V. parahaemolyticus* (153, 154). In the following years, it was shown that the T3SS of strain AM-19226 was functional and required for intestinal colonization of infant mice (155). Animals infected with the strain developed severe diarrhea and showed disruption of the small intestinal epithelium, contrary to what is observed in rabbits infected with pandemic O1 *V. cholerae* strains (156). Interestingly, Mauve alignment (157) of the T3SS-encoding islands of the here-described environmental *V. cholerae* strains to the genome of strain AM-19226 (contigs AATY02000003 and AATY02000004) demonstrated that the islands are remarkably similar, globally sharing the same gene content and organization with few exceptions. Furthermore, we performed an in-depth manual analysis by genome comparisons and BLAST search of all genes inside the predicted T3SS cluster and their gene products. By doing so, we were able to identify genes that encode all main components of the T3SS system (Figure 3.5b). We display the identified genes with two kind of nomenclatures: the original species-specific name given by the closest hit from the BLAST search, and the unified secretion and cellular translocation nomenclature Sct (158).

Eleven T3SS effectors were previously identified in strain AM-19226 based on their ability to inhibit growth when heterologously expressed in yeast or when translocated into HeLa cells *in vitro* (159). Out of these eleven *bona fide* effectors, only two were apparently

absent from the T3SS clusters of the environmental strains (*vopW* and *vopZ)*. Furthermore, out of those eleven effectors identified in AM-19226, three have been experimentally studied and validated, namely VopF, VopE, and VopX. VopF has formin homology 1-like (FH1-like) and WASP homology 2 (WH2) domains, and it is implicated in changes in the proper function of actin polymerization, which is necessary for adequate infant mice intestinal colonization by T3SS+ *V. cholerae* (155). VopE, a homolog of the *Yersinia* YopE protein (160), contains a GTPase-activating protein (GAP) domain and localizes to mitochondria during infection, disturbing the activity of mitochondrial Rho GTPases, which, ultimately, interferes with the immune response (161, 162). VopX of strain AM-19226 is important for infant mice colonization and was suggested to communicate with the mitogen-activated protein kinase (MAPK) pathway, involved in cell wall integrity (159).

Finally, we also identified two homologs of the ToxR-like virulence regulators $VttR_A$ and $VttR_B$ of strain AM-19226 (110). These regulators are involved in bile acid-induced expression of the T3SS in AM-19226 (163). Interestingly, $vttR_A$ was initially annotated by NCBI's PGAP as *hilA*, which encodes the main transcriptional regulator of one of the best studied T3SSs encoded on the *Salmonella* pathogenicity island I (SPI-1) (164). These similarities strongly support the notion that these strains encode a *bona fide* T3SS. Future functionality studies will elucidate their relevance in the biology of these environmental *V. cholerae* strains.

Expression profile

The transcript levels of genes within this island were evaluated by RNA sequencing of the strains SA5Y (clade B) and SL5Y (clade C). In both strains, the *nan-nag* region was expressed at higher levels at low cell density compared to the low cell density situation, which was also the case for the *nan-nag* cluster on VPI-2 of the pandemic strain A1552 (data not shown). Related to the T3SS, almost all genes were expressed at low levels during

both growth phases. Interestingly, the gene encoding the transcriptional regulator $VttR_A$ was highly expressed in both strains, especially at low cell density, while the gene encoding the other regulator, $VttR_B$, seems barely expressed.

Experimental outlook

As the transcript levels of the T3SS genes were low for both strains, we will first genetically engineer strains carrying inducible copies of $vttR_A$ and/or $vttR_B$. Upon induction, the expression of the T3SS genes will be checked by qRT-PCR. We will also test the production of FLAG-tagged VscN2 by Western blot analysis. If expression can be triggered by artificially induction of one or both of the regulatory genes, we will establish a secretion assay to check the functionality of the T3SS, as previously described for *V. parahemolyticus* (165) and *Salmonella* (166) for example. We will then test variants that lack, for instance, the ATP synthase from the sorting platform YscN/SctN or core structural components. Finally, we will generate mutant strains lacking selected T3SS-related gene(s) and compare their toxicity against Caco2 (167) and HeLa cells (168), or grazing amoebae (such as *D. discoideum* or *Acanthamoeba castellanii*; see chapters 2 and 6), to the toxicity of the parental WT strain.

3.3 CONCLUSION

The current ongoing project has been an exploratory expedition into the diversity of horizontally acquired genomic islands in environmental isolates of *V. cholerae.* The *in silico* obtained results are very promising, as we uncovered a remarkable diversity of MGEs that might play important roles in *V. cholerae*'s environmental lifestyle. Importantly, our transcriptomic data suggests that the majority of these islands is expressed, even under standard laboratory conditions. Future experiments using the above-described approaches will assess their significance for these bacteria.

3.4 MATERIAL AND METHODS

3.4.1 Bacterial strains

V. cholerae strains used in this study are listed in Supplementary table 3.1.

3.4.2 *In silico* characterization of environmental genomic islands

The two closed chromosomes (GenBank accession numbers can be found in Supplementary table 3.1) of each of the fifteen environmental isolates (43, 47) were manually inspected for genomic islands that are absent from the genome of a representative 7th pandemic O1 El Tor strain (A1552; (47)). Each chromosome sequence of the environmental isolates was aligned to the corresponding chromosome of A1552 using the Mauve algorithm (157) in Geneious 10.2.6 (https://www.geneious.com).

First, we inspected whether the environmental strains harbored any genomic islands at the same genomic location where the prominent pathogenicity islands VPI-1, VPI-2, VSP-I and VSP-II are located in pandemic strains. Next, we examined genomic islands found in other locations on chromosomes 1 and 2. The predicted function of all genes was initially annotated by the Genomic Technology Facility (GTF) of the University of Lausanne using PROKKA. Reannotation of the majority of genomes occurred upon submission of the sequencing data to NCBI (using the Prokaryotic Genome Annotation Pipeline, PGAP) (43, 47). Additional manual ORF translation was done using Geneious 10.2.6 followed by BLASTp. For the sake of keeping the extend of this study manageable, we focused our work on the investigation of islands that seemed most diverse regarding their genomic content.

3.4.3 RNA sequencing

Overnight cultures were back-diluted 1:100 in LB medium and grown with agitation at 30 °C for 2 h or 6 h (OD_{600} at the time of harvesting can be found in Supplementary table 3.3).

Cells were harvested and washed with Phosphate Buffered Saline (PBS pH 7.4; Life Technologies) at 4 °C, followed by lysis with Tri reagent (Sigma-Aldrich). RNA preparation and DNase treatment were performed as previously described (169). After DNAse treatment, an additional purification step was performed using the RNeasy MinElute Cleanup Kit (Qiagen).

Sample processing, sequencing, and bioinformatic analyses were performed by the company Microsynth (Switzerland). Illumina's TruSeq Stranded Total RNA Library Prep Gold kit including ribodepletion was used to construct libraries using the total RNA as template. Subsequently, the Illumina NextSeq 550 platform and a high output v2.5 kit (single-end 1x75bp) were used to sequence the libraries. The final single-end reads that passed the Illumina's chastity filter were subject to de-multiplexing and trimming of Illumina adaptor residuals using Illumina's bcl2fastq software version v2.20.0.422 (no further refinement or selection). Quality of the reads in fastq format was checked with the software FastQC (version 0.11.8) (http://www.bioinformatics.babraham.ac.uk/projects/fastqc/). The reads were mapped to the reference genomes via bowtie2 (version 2.3.5.1) (170) in local mapping mode with very sensitive pre-settings. To count the uniquely mapped reads to annotated genes, the software htseq-count (HTSeq version 0.11.2) (171) was used. Normalization of the raw counts and differential gene expression analysis was carried out with help of the R software package DESeq2 (version 1.22.2) (172). Visualization of sample gene expression profiles was realized with the R software package Rtsne (version 0.15) (173).

3.5 FIGURES

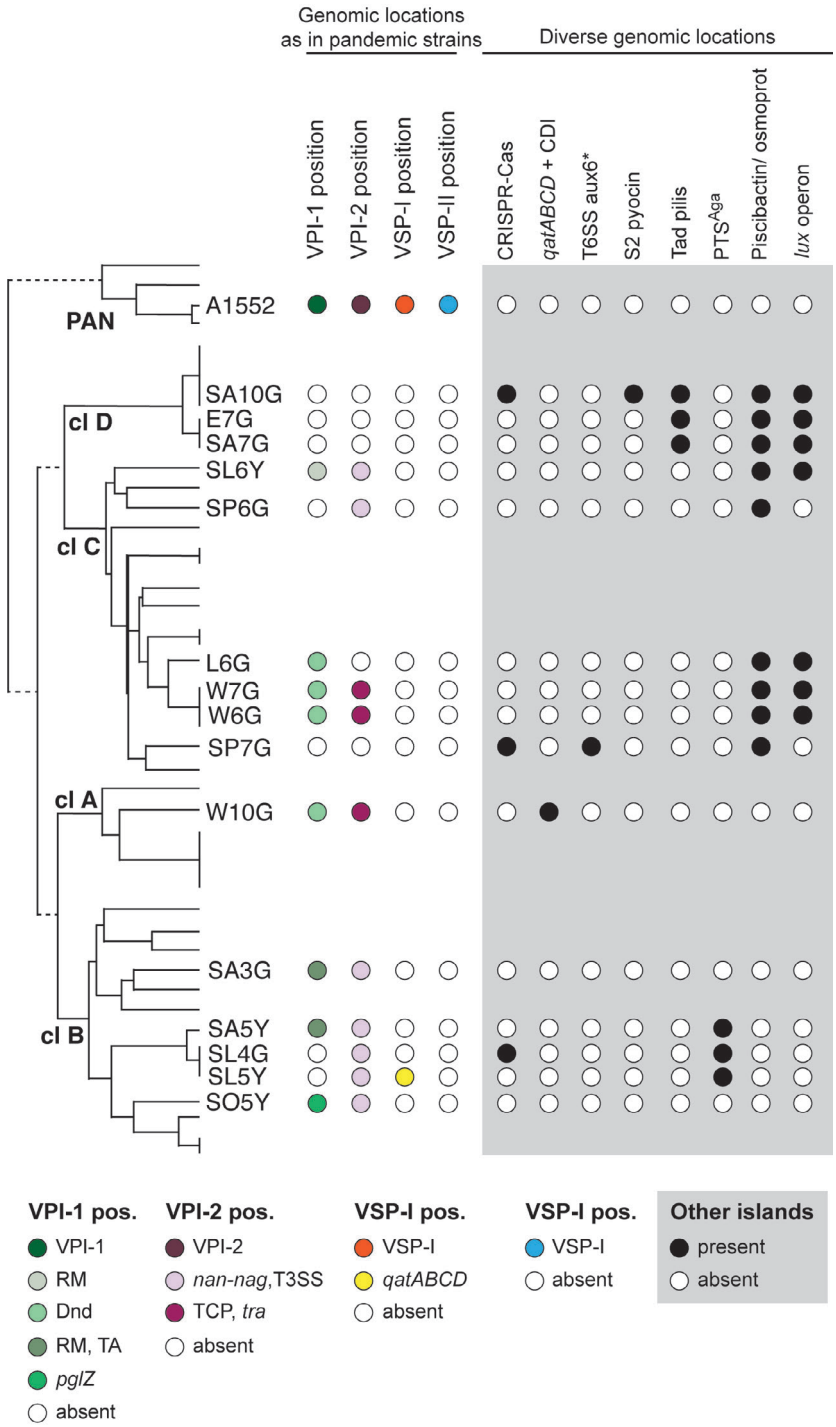


Figure 3.1 Genomic islands found in environmental isolates of *V. cholerae.*

Phylogenetic tree freely adapted from ref (44) [not to scale]. Environmental genomic islands found in the same locations as VPI-1, VPI-2, VSP-I and VSP-II in pandemic *V. cholerae* (A1552; shown on top) or in other chromosomal locations are depicted. Their genetic content is either color-coded or scored for their presence (black circles) or absence (white circles) in the grey box on the right. RM: restriction-modification system; Dnd: Dnd defense system; TA: toxin-antitoxin module; *pglZ*: part of phage growth limitation system; TCP: homologs of *tcp* (toxin co-regulated pilus) genes; *tra*: genes related to conjugation F-pilus; *qatABCD*: antiphage defense system. *The T6SS auxiliary cluster 6 was described in (43).

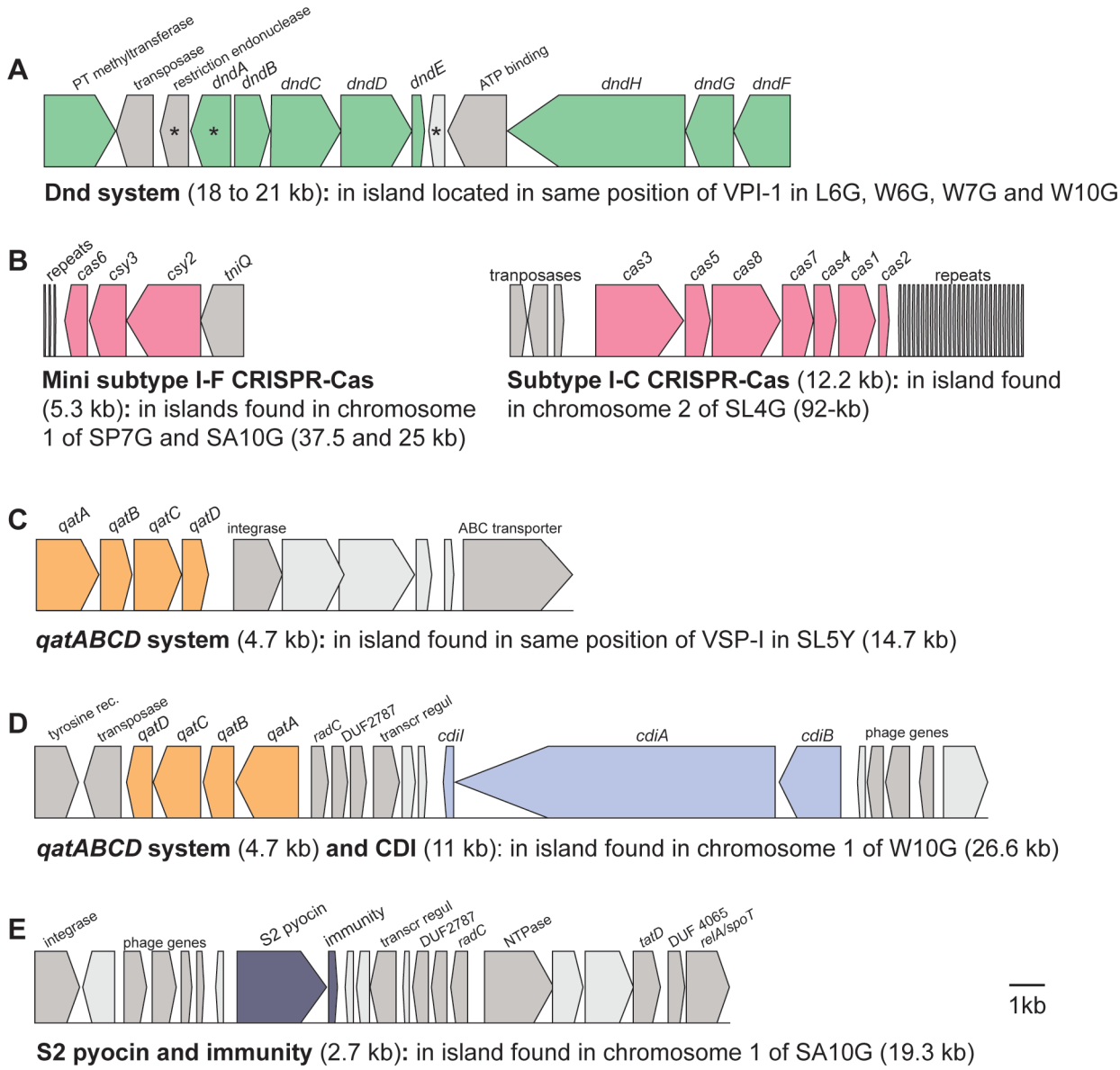


Figure 3.2 Antiphage and antibacterial defense systems found in genomic islands of environmental strains.

(A) Dnd defense system found in strains L6G, W6G, W7G and W10G. The transposase gene was only found in the island of W10G. *are genes not found in the island of strain L6G. **(B)** Different CRISPR-Cas modules found in environmental strains. The mini I-F module (left) was found on chromosome 1 islands of strains SP7G (4 spacer/repeats - S/Rs) and SA10G (3 S/Rs). The I-C module (28 S/Rs) (right) was found on an island in chromosome 2 of strain SL4G. **(C)** A *qatABCD* antiphage defense system of strain SL5Y was found in the same location as VSP-I in pandemic *V. cholerae* strains. The same operon was found in an island on chromosome 1 of strain W10G **(D)**, along with a contact-dependent inhibition (CDI) system. **(E)** Genes encoding a S2 pyocin-like protein and its immunity protein were found in an island on chromosome 1 of strain SA10G. Dark grey arrows represent genes that could be identified by BLAST but that are (in theory) not related to the main focus given to the island. Light grey arrows represent hypothetical genes.

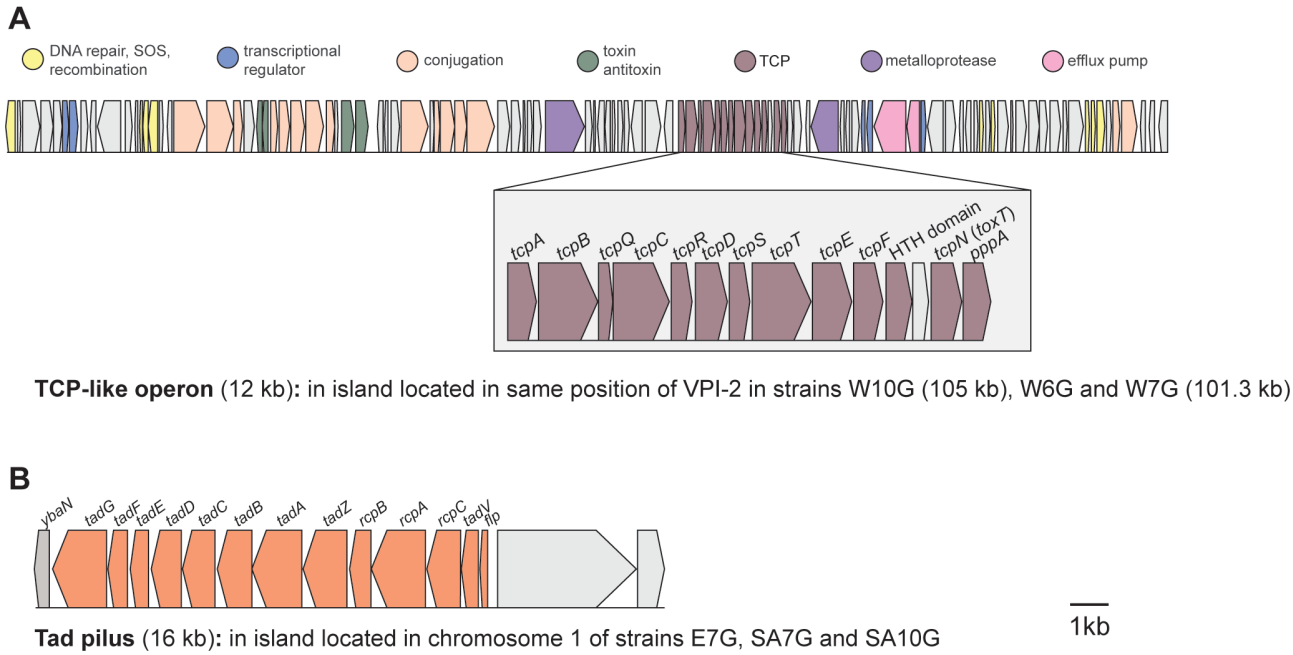


Figure 3.3 Clusters for the production of type IV pili (T4P) found in genomic islands of environmental strains.

(A) Large genomic island (101-105kb) in strains W6G, W7G, and W10G. Predicted functions are color-coded. Details of the *tcp* homologous operon are enlarged in the grey box. These islands in the environmental strains are located at the same genomic site as VPI-2 in 7th pandemic strains. **(B)** A set of *tad* genes (encoding a Tad pilus) was found on a genomic island in clade D environmental strains. Dark grey arrows represent genes that could be identified by BLAST but that are (in theory) not related to the main focus given to the island. Light grey arrows represent hypothetical genes.

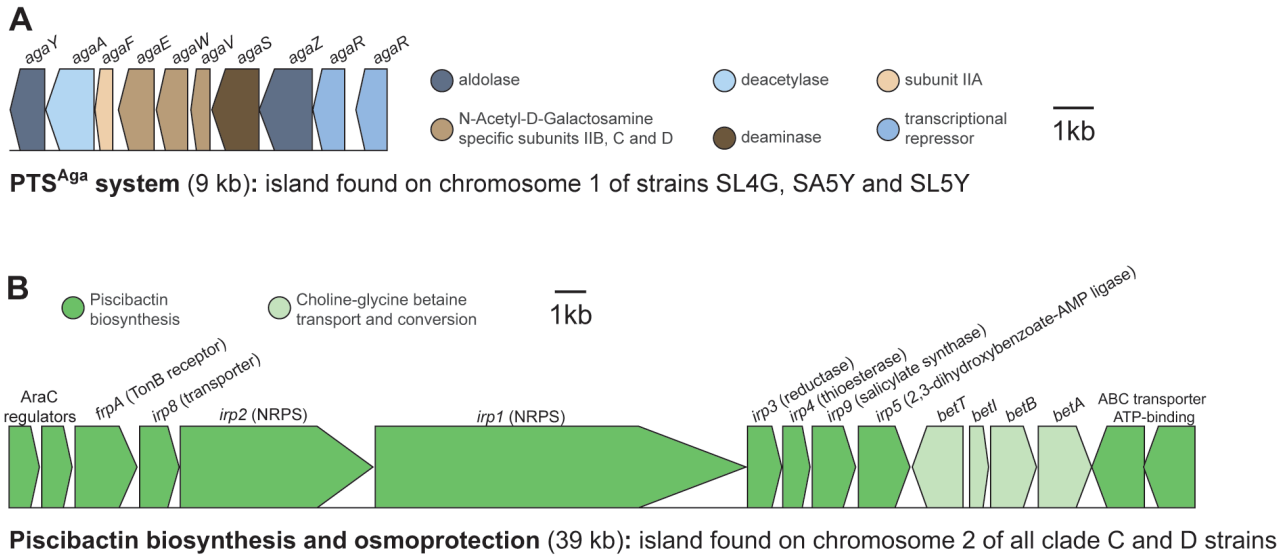


Figure 3.4 Environmental genomic islands encoding metabolism-related genes.

(A) A 9 kb island with genes encoding subunits for N-Acetylgalactosamine (Aga) utilization is found on chromosome 1 of strains SL4G, SA5Y and SL5Y. **(B)** A 39 kb island is widespread in the collection of environmental *V. cholerae* isolates, where it is present in all clade C and D strains. This island contains a piscibactin siderophore biosynthesis cluster and a small operon of *bet* genes involved in the transport and conversion of choline resulting in the osmolyte glycine betaine.

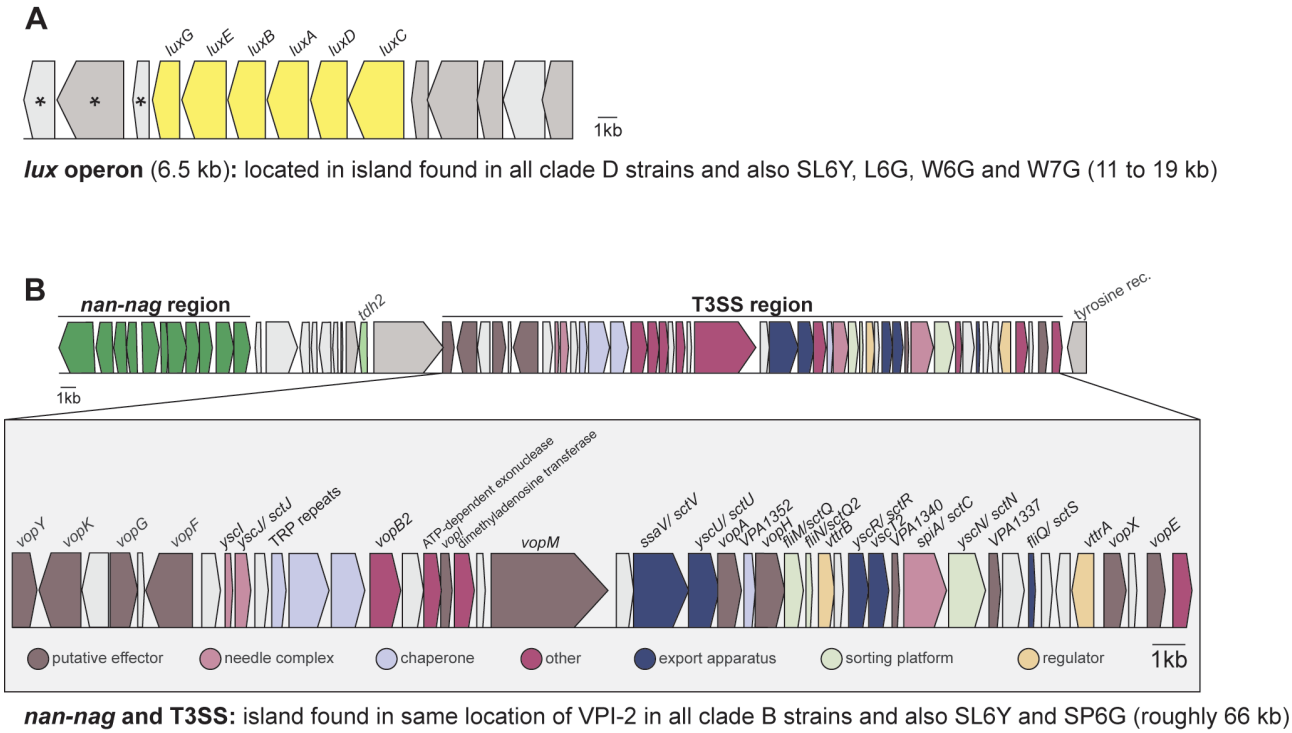
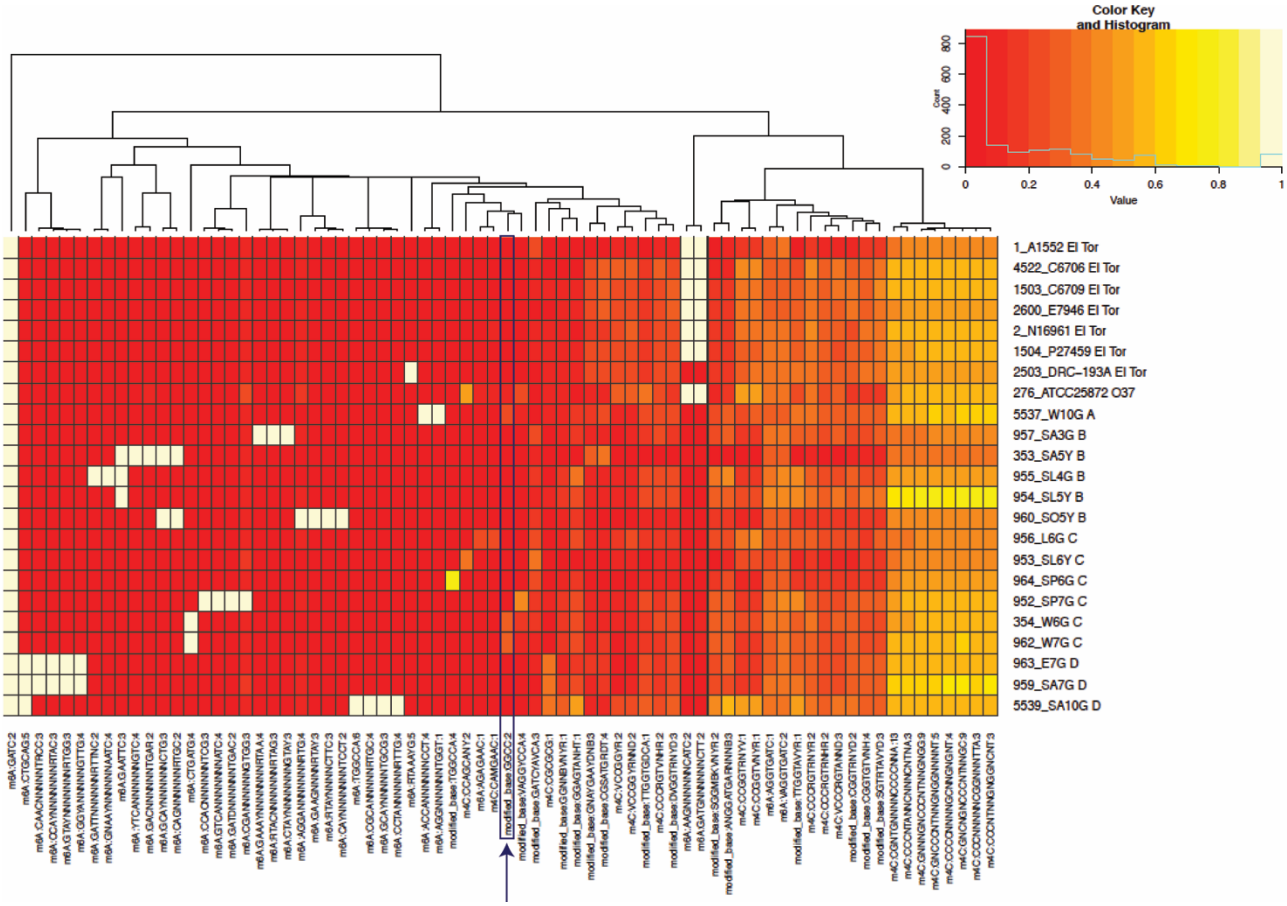


Figure 3.5 Gene clusters relevant for interactions with diverse host(s) found on genomic islands of environmental strains.

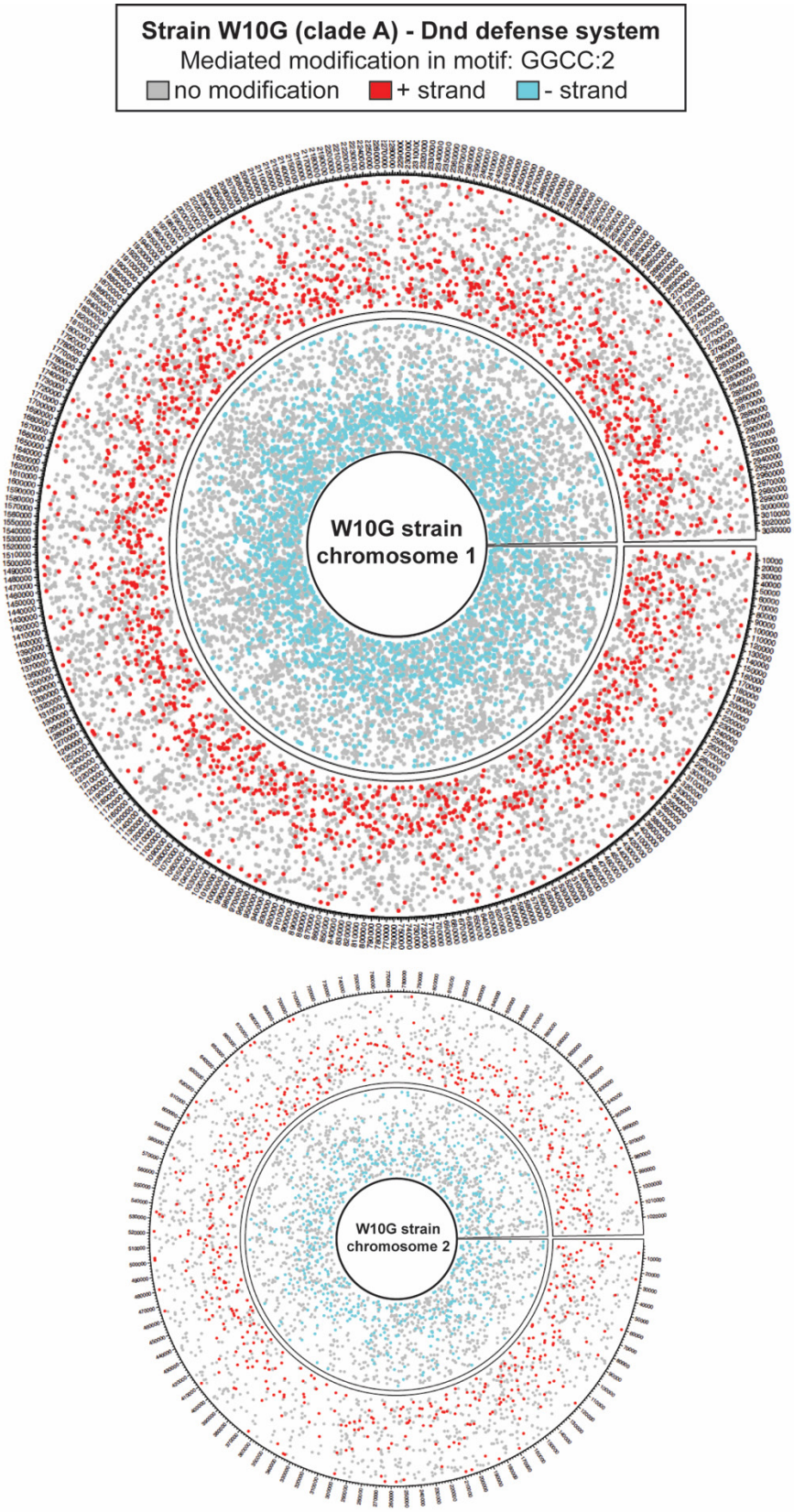
(A) The 6.5 kb *luxCDABEG* operon for bioluminescence was found on genomic islands (11 to 19-kb long) of clade D strains and also strains SL6Y, L6G, W6G and W7G (clade C). Genes marked with an asterisk symbolize genes that varied among strains (see text for details). Strain SA10G contains three extra ORFS at the extremity of the cluster, which are not represented in the scheme. **(B)** A roughly 66-kb island was found in the same location as VPI-2 in pandemic strains in all clade B strains and also strains SL6Y and SP6G (clade C). The island consists of a *nan-nag* region for sialic acid utilization (similar to the *nan-nag* region in VPI-2). The rest of the island contains genes for a type III secretion system (T3SS), which are color-coded according to their predicted role. Dark grey arrows represent genes that could be identified by BLAST but that are (in theory) not related to the main focus given to the island. Light grey arrows represent hypothetical genes.

3.6 SUPPLEMENTARY FIGURES



Supplementary figure 3.1 Heatmap of modified genomic motifs.

DNA motifs and modifications found in at least 25% of the genome of one sample with the corresponding modification based on SMRT sequencing of pandemic and environmental *V. cholerae* strains. Color code indicates the proportion of the motifs that were modified, from 0 (red) to 1 (light yellow). Putative mark derived from Dnd defense system is indicated with an arrow. Genomes of strains W10G, W6G and W7G had approximately 26% of these motifs modified.



Supplementary figure 3.2 Genomic landscape of putative Dnd-mediated modification in strain W10G.

Motif GGCC found unmodified (grey), modified in the + strand (red) or the – strand (blue) in the chromosome 1 and 2 of strain W10G (clade A).

3.7 SUPPLEMENTARY TABLES

Supplementary table 3.1 *V. cholerae* strains used in this study.

Strain names	Genotype, description and genome accession numbers	Internal strain number	Reference
A1552	Wild-type, O1 El Tor Inaba; Rif$^{R.}$ WGS GenBank accession numbers CP028894 (chr1) and CP028895 (chr2)*	MB_1	(173)
W10G	Environmental isolate (clade A) collected in Waddell Creek (CA, USA) in October 2004. WGS GenBank accession numbers CP053794 (chr1) and CP053795 (chr2)**	MB_5537	(44)
SA3G	Environmental isolate (clade B) collected in Old Salinas River (CA, USA) in March 2004. WGS GenBank accession numbers CP053744 (chr1) and CP053745 (chr2)**	MB_957	(44)
SA5Y	Environmental isolate (clade B) collected in Old Salinas River (CA, USA) in May 2004. WGS GenBank accession numbers CP028892 (chr1) and CP028893 (chr2)*	MB_353	(44)
SL4G	Environmental isolate (clade B) collected in San Lorenzo River (CA, USA) in April 2004. WGS GenBank accession numbers CP053796 (chr1) and CP053797 (chr2)**	MB_955	(44)
SL5Y	Environmental isolate (clade B) collected in San Lorenzo River (CA, USA) in May 2004. WGS GenBank accession numbers CP053798 (chr1) CP053799 and (chr2)**	MB_954	(44)
SO5Y	Environmental isolate (clade B) collected in Soquel Creek (CA, USA) in May 2004. WGS GenBank accession numbers CP053800 (chr1) and CP053801 (chr2)**	MB_960	(44)
L6G	Environmental isolate (clade C) collected in Lagunitas Creek (CA, USA) in June 2004. WGS GenBank accession numbers CP053802 (chr1) CP053803 (chr2)**	MB_956	(44)
SL6Y	Environmental isolate (clade C) collected in San Lorenzo River (CA, USA) in June 2004. WGS GenBank accession numbers CP053804 (chr1) and CP053805 (chr2)**	MB_953	(44)
SP6G	Environmental isolate (clade C) collected in San Pedro Creek (CA, USA) in June 2004. WGS GenBank accession numbers CP053806 (chr1) and CP053807 (chr2)**	MB_964	(44)
SP7G	Environmental isolate (clade C) collected in San Pedro Creek (CA, USA) in July 2004. WGS GenBank accession numbers CP053808 (chr1) and CP053809 (chr2)**	MB_952	(44)
W6G	Environmental isolate (clade C) collected in Waddell Creek (CA, USA) in June 2004. WGS GenBank accession numbers CP053810 (chr1) and CP053811 (chr2)**	MB_354	(44)
W7G	Environmental isolate (clade C) collected in Waddell Creek (CA, USA) in July 2004. WGS GenBank accession numbers CP053813 (chr1) and CP053814 (chr2)**	MB_962	(44)
E7G	Environmental isolate (clade D) collected in Moss Landing Harbor (CA, USA) in July 2004. WGS GenBank accession numbers CP053822 (chr1) and CP053823 (chr2)**	MB_963	(44)
SA7G	Environmental isolate (clade D) collected in Old Salinas River (CA, USA) in July 2004. WGS GenBank accession numbers CP053816 (chr1) CP053817 and (chr2)**	MB_959	(44)
SA10G	Environmental isolate (clade D) collected in Old Salinas River (CA, USA) in October 2004. WGS GenBank accession numbers CP053820 (chr1) and CP053821 (chr2)**	MB_5539	(44)

* closed genomes published in GenBank with publication (47).

** closed genomes published in GenBank with publication (43).

Supplementary table 3.2 Locus tags of operons located on genomic islands described in this study.

Operons	Locus tags
Dnd system	**W10G:** HPY12_10380 - HPY12_10440 **L6G:** HPY16_09335 - HPY16_09375 **W6G:** HPY05_10065 - HPY05_10120 **W7G:** HPY09_10070 - HPY09_10125
CRISPR-Cas	**SP7G (chr1):** HPY08_08365 - HPY08_08465 (full island); HPY08_08395 - HPY08_08410 (CRISPR-Cas operon, including transposase) **SA10G (chr1):** HPY11_08770 - HPY11_08865 (full island); HPY11_08795 - HPY11_08810 (CRISPR-Cas operon, including transposase) **SL4G (chr2):** HPY13_16695 - HPY13_17120 (full island); HPY13_16905 - HPY13_16950 (CRISPR-Cas operon, including transposases)
***qatABCD* system**	**W10G:** HPY12_09495 - HPY12_09590 (full island); HPY12_09520 - HPY12_09505 (*qatABCD* operon) **SL5Y:** HPY14_13080 - HPY14_13125 (full island); HPY14_13080 - HPY14_13095 (*qatABCD* operon)
CDI system (*cdiBAI*)	**W10G:** HPY12_09495 - HPY12_09590 (full island; same as *qatABCD*); HPY12_09565 - HPY12_09555 (*cdiBAI*)
S2 pyocin & immunity	**SA10G:** HPY11_00665 - HPY11_00770 (full island); HPY11_00700 - HPY11_00705 (pyocin + immunity)
TCP-like pilus	**W10G:** HPY12_05275- HPY12_05815 (full island); HPY12_05555 - HPY12_05620 (TCP operon) **W6G:** HPY05_05255 - HPY05_05770 (full island); HPY05_05525 – HPY05_05590 (TCP operon) **W7G:** HPY09_05255 - HPY09_05770 (full island); HPY09_05525 - HPY09_05590 (TCP operon)
Tad-type pilus	**E7G:** HPY10_13170 - HPY10_13245 **SA7G:** HPY17_13130 - HPY17_13205 **SA10G:** HPY11_13220 - HPY11_13295
N-Acetyl galactosamine PTS system	**SA5Y:** Sa5Y_VC01785 - Sa5Y_VC01794 **SL4G:** HPY13_09020 - HPY13_09065 **SL5Y:** HPY14_08775 - HPY14_08820
Piscibactin production and osmoprotection	**L6G:** HPY16_16905 - HPY16_16980 **SL6Y:** HPY06_16855 - HPY06_16930 **SP6G:** HPY07_17225 - HPY07_17300 **SP7G:** HPY08_17085 - HPY08_17160 **W6G:** HPY05_17680 - HPY05_17785 **W7G:** HPY09_17675 - HPY09_17780 **E7G:** HPY10_17830 - HPY10_17905 **SA7G:** HPY17_17590 - HPY17_17665 **SA10G:** HPY11_17500 - HPY11_17575
***lux* operon**	**L6G:** HPY16_06145 - HPY16_06200 (full island); HPY16_06150 - HPY16_06175 (*lux* operon) **SL6Y:** HPY06_06495 - HPY06_06550 (full island); HPY06_06500 - HPY06_06525 (*lux* operon) **W6G:** HPY05_06740 - HPY05_06795 (full island); HPY05_06745 - HPY05_06770 (*lux* operon) **W7G:** HPY09_06740 - HPY09_06795 (full island); HPY09_06745 - HPY09_06770 (*lux* operon) **E7G:** HPY10_06410 - HPY10_06475 (full island); HPY10_06425 - HPY10_06450 (*lux* operon) **SA7G:** HPY17_06395 - HPY17_06460 (full island); HPY17_06410 - HPY17_06435 (*lux* operon) **SA10G:** HPY11_06400 - HPY11_06485 (full island); HPY11_06435 - HPY11_06460 (*lux* operon)
***nan-nag* region and T3SS**	**SA3G:** HPY04_05325 - HPY04_05650 **SA5Y:** Sa5Y_VC01069 - Sa5Y_VC01135 **SL4G:** HPY13_05235 - HPY13_05560 **SL5Y:** HPY14_05240 - HPY14_05560 **SO5Y:** HPY15_05250 - HPY15_05590 **SL6Y:** HPY06_05260 - HPY06_05585 **SP6G:** HPY07_05285 - HPY07_05650

Supplementary table 3.3 Environmental strains used for RNAseq and their respective OD_{600} at the time of RNA harvesting.

		2 h growth			6 h growth		
Clade	**Strain**	**Repl 1**	**Repl 2**	**Repl 3**	**Repl 1**	**Repl 2**	**Repl 3**
A	W10G	0.318	0.279	0.221	5.026	3.974	4.084
B	SA5Y	0.543	0.333	0.501	3.933	3.669	4.032
	SL5Y	0.653	0.593	0.649	3.937	3.480	3.847
C	L6G	0.352	0.300	0.402	4.684	4.769	4.390
	SP7G	0.576	0.340	0.414	4.238	4.361	4.273
D	SA10G	0.794	0.639	0.647	6.444	4.627	4.587

Supplementary table 3.4 Expression data (normalized counts) for Dnd defense system. RNAseq data exemplified for strains W10G and L6G.

		2 h growth				6 h growth				Ratio
Gene	**Gene product**	**Repl 1**	**Repl 2**	**Repl 3**	**Average**	**Repl 1**	**Repl 2**	**Repl 3**	**Average**	**6h/ 2h**
Strain W10G										
-	PT-Methyltransferase	1935.3	1387.8	1323.0	1548.7	1831.8	2168.8	1772.2	1924.2	1.2
-	Restriction endonuclease	2842.2	2951.6	2251.6	2681.8	1200.6	1372.6	1401.5	1324.9	0.5
dndA	Cysteine desulfurase	3898.1	4417.5	3589.0	3968.2	1420.0	1436.0	1474.7	1443.6	0.4
dndB	Sulfur modification protein	4649.0	6589.8	6361.0	5866.6	933.4	953.6	932.2	939.7	0.2
dndC	Sulfurtransferase	8007.7	8719.5	8122.8	8283.3	1247.7	1493.6	1356.5	1365.9	0.2
dndD	Sulfur modification protein	10428.9	10476.6	9300.0	10068.5	1255.4	1455.3	1437.0	1382.6	0.1
dndE	Sulfur modification protein	2021.6	2049.2	1955.1	2008.6	260.3	265.8	323.4	283.2	0.1
-	Hypothetical	11401.1	8784.2	5384.4	8523.2	1191.1	1449.8	1679.8	1440.2	0.2
-	ATP binding protein	34859.3	21866.3	16450.7	24392.1	10516.9	13368.6	13729.9	12538.5	0.5
dptH	PT-dependent restriction protein	28805.3	14321.6	15170.1	19432.4	16144.1	20459.1	17554.7	18052.6	0.9
dptG	PT-dependent restriction protein	8892.5	3933.1	3757.5	5527.7	2949.1	3743.2	3442.0	3378.1	0.6
dptF	PT-dependent restriction protein	12068.1	6521.8	5449.6	8013.1	3626.3	4181.4	4204.2	4004.0	0.5
Strain L6G										
	PT-Methyltransferase	3647.1	3050.7	3794.0	3497.2	5151.5	4809.5	4827.1	4929.4	1.4
dndB	Sulfur modification protein	1633.2	1607.3	1489.2	1576.6	631.8	1338.0	942.8	970.9	0.6
dndC	Sulfurtransferase	2504.2	2124.1	1890.7	2173.0	706.2	1276.1	1010.2	997.5	0.5
dndD	Sulfur modification protein	5269.2	4761.1	4853.6	4961.3	2358.0	3451.0	2631.0	2813.3	0.6
dndE	Sulfur modification protein	1406.5	1661.1	1609.6	1559.1	271.6	597.9	387.0	418.9	0.3
-	DUF262 domain-containing protein	20275.9	17261.5	16573.1	18036.8	4729.7	9972.5	9423.4	8041.9	0.4
dptH	PT-dependent restriction protein	15221.6	13481.3	15607.2	14770.0	13851.2	15680.7	14695.2	14742.4	1.0
dptG	PT-dependent restriction protein	4700.1	4095.0	4807.4	4534.2	2230.7	3183.1	2824.5	2746.1	0.6
dptF	PT-dependent restriction protein	7129.1	6493.0	7046.4	6889.5	2249.9	4074.4	3405.3	3243.2	0.5

Supplementary table 3.5 Expression data (normalized counts) for CRISPR-Cas modules found in strains SP7G and SA10G.

		2 h growth				6 h growth				Ratio
Gene	**Gene product**	**Repl 1**	**Repl 2**	**Repl 3**	**Average**	**Repl 1**	**Repl 2**	**Repl 3**	**Average**	**6h/ 2h**
Strain SP7G										
cas6f	Type I-F CRISPR endoribonuclease Cas6/Csy4	702.4	416.9	402.5	507.3	8359.6	7575.3	6765.1	7566.7	14.9
csy3	Type I-F CRISPR protein Csy3	1182.3	594.2	623.3	799.9	14170.0	13693.1	12197.0	13353.4	16.7
csy2	CRISPR- protein Csy2	1401.2	998.9	1107.8	1169.3	6572.9	5997.4	5629.7	6066.7	5.2
tniQ	TniQ family protein	519.3	491.7	642.8	551.3	1440.8	1371.2	1026.2	1279.4	2.3
Strain SA10G										
cas6f	Endoribonuclease Cas6/Csy4	690.6	633.5	424.9	583.0	6252.3	5225.3	5037.3	5505.0	9.4
csy3	CRISPR- protein Csy3	1136.9	840.6	736.2	904.6	15762.9	14387.3	11667.9	13939.3	15.4
csy2	CRISPR- protein Csy2	1469.6	1863.1	1415.4	1582.7	6442.4	5651.8	5072.5	5722.2	3.6
tniQ	TniQ family protein	537.8	777.1	602.5	639.1	1243.9	1167.9	805.5	1072.4	1.7

Supplementary table 3.6 Expression data (normalized counts) for *qatABCD* defense system found in strains W10G and SL5Y.

		2 h growth				6 h growth				Ratio
Gene	**Gene product**	**Repl 1**	**Repl 2**	**Repl 3**	**Average**	**Repl 1**	**Repl 2**	**Repl 3**	**Average**	**6h/ 2h**
Strain W10G										
qatA	ATPase	4893.9	1969.6	1874.7	2912.7	3361.1	4187.3	3746.8	3765.1	1.3
qatB	Hypothetical	3266.2	1399.3	1191.7	1952.4	2919.8	3402.8	3177.6	3166.8	1.6
***qatC* (*queC*)**	7-cyano-7-deazaguanine synthase	6700.9	3748.8	3446.4	4632.0	4589.6	5805.2	4953.8	5116.2	1.1
***qatD* (*tatD*)**	TatD family deoxyribonuclease	2683.3	1630.7	1499.0	1937.7	2979.2	3658.4	3272.7	3303.4	1.7
Strain SL5Y										
qatA	ATPase	6367.9	3833.9	3067.2	4423.0	13851.9	14832.0	13468.4	14050.8	3.2
qatB	Hypothetical	2795.2	1088.6	1057.7	1647.2	7059.3	7097.0	6489.0	6881.8	4.2
***qatC* (*queC*)**	7-cyano-7-deazaguanine synthase	5613.9	3407.1	2602.5	3874.5	13939.0	14555.9	12400.6	13631.8	3.5
***qatD* (*tatD*)**	TatD family deoxyribonuclease	2085.3	1898.9	1544.8	1843.0	3756.3	3585.7	3144.5	3495.5	1.9

Supplementary table 3.7 Expression data (normalized counts) for contact-dependent inhibition (CDI) system found in strain W10G.

		2 h growth				6 h growth				Ratio
Gene	**Gene product**	**Repl 1**	**Repl 2**	**Repl 3**	**Average**	**Repl 1**	**Repl 2**	**Repl 3**	**Average**	**6h/ 2h**
cdiI	Hypothetical	3241.8	2311.0	1644.8	2399.2	1300.7	1494.9	1388.6	1394.7	0.6
cdiA	Filamentous hemagglutinin N-terminal domain-containing protein	82107.3	62968.0	54347.4	66474.2	24864.7	28025.5	24288.5	25726.3	0.4
cdiB	ShlB/FhaC/HecB family hemolysin secretion/activation protein	20096.1	20509.6	15609.2	18738.3	4120.8	4171.5	4208.2	4166.8	0.2

Supplementary table 3.8 Expression data (normalized counts) for S2-type pyocin system found in strain SA10G.

		2 h growth				6 h growth				Ratio
Gene	**Gene product**	**Repl 1**	**Repl 2**	**Repl 3**	**Average**	**Repl 1**	**Repl 2**	**Repl 3**	**Average**	**6h/ 2h**
psy2	S-type pyocin domain-containing protein	4508.7	3299.9	3198.2	3668.9	7784.8	7089.2	7322.8	7398.9	2.0
-	Bacteriocin immunity protein	509.9	1184.4	921.5	871.9	1017.9	1114.2	1032.2	1054.8	1.2

Supplementary table 3.9 Expression data (normalized counts) for genes encoding TCP-like pilus. RNAseq data exemplified for strain W10G.

		2 h growth				6 h growth				Ratio
Gene	**Gene product**	**Repl 1**	**Repl 2**	**Repl 3**	**Average**	**Repl 1**	**Repl 2**	**Repl 3**	**Average**	**6h/ 2h**
tcpA	Type IV pilus major pilin	330.1	304.6	237.4	290.7	223.5	260.6	231.6	238.5	0.8
tcpB	Type II secretion system protein	213.4	218.8	190.6	207.6	41.6	37.6	40.9	40.1	0.2
tcpQ	Pilus assembly protein	63.1	58.2	69.0	63.5	16.1	7.1	16.0	13.0	0.2
tcpC	Type II and III secretion system protein	583.9	442.6	396.3	474.3	285.6	312.3	362.5	320.1	0.7
tcpR	Hypothetical	185.3	171.0	138.8	165.0	62.2	77.2	81.5	73.6	0.4
tcpD	Pilus assembly protein	306.3	252.8	215.4	258.2	64.1	103.7	72.8	80.2	0.3
tcpS	Hypothetical	200.6	172.9	181.4	185.0	165.5	133.0	159.0	152.5	0.8
tadA* (*tcpT*)	Flp pilus assembly complex ATPase component TadA	1275.4	1167.7	1029.1	1157.4	1820.8	2037.0	1835.5	1897.8	1.6
tcpE	Type II secretion protein F	945.9	912.7	736.0	864.9	669.4	788.3	745.2	734.3	0.8
-	Prepilin peptidase	590.6	631.8	407.6	543.3	139.1	168.4	144.1	150.5	0.3
-	Helix-turn-helix domain-containing protein	216.8	299.7	165.7	227.4	5.3	8.6	15.9	9.9	0.0
-	Hypothetical	71.8	78.5	75.6	75.3	12.3	7.5	12.7	10.8	0.1
toxT* (*tcpN*)	Helix-turn-helix Transcriptional regulator	93.0	133.6	92.9	106.5	118.1	131.9	135.2	128.4	1.2
-	Hypothetical	50.2	51.1	41.4	47.6	35.8	33.2	19.9	29.7	0.6

*Automatic annotation gene name (with corresponding gene product description in next column), while gene name in brackets correspond to nomenclature most commonly used in the operon context.

Supplementary table 3.10 Expression data (normalized counts) for genes encoding a Tad-type pilus. RNAseq data exemplified for SA10G.

		2 h growth				6 h growth				Ratio
Gene	**Gene product**	**Repl 1**	**Repl 2**	**Repl 3**	**Average**	**Repl 1**	**Repl 2**	**Repl 3**	**Average**	**6h/ 2h**
tadG	VWA domain-containing protein	626.8	680.8	628.0	645.2	1388.1	1122.1	1118.3	1209.5	1.9
tadF	Hypothetical	173.3	253.1	230.1	218.8	147.8	181.6	164.0	164.5	0.8
tadE	Pilus assembly protein	162.1	155.2	183.7	167.0	160.2	166.0	156.1	160.8	1.0
tadD	Tetratricopeptide repeat protein	287.8	314.1	408.3	336.8	475.2	421.1	397.7	431.4	1.3
tadC	Type II secretion system F family protein	263.6	214.5	268.2	248.8	242.0	225.3	269.1	245.5	1.0
tadB	Pilus assembly protein	210.4	212.3	230.4	217.7	261.8	259.0	234.5	251.7	1.2
tadA	CpaF family protein	296.0	216.3	269.1	260.5	509.6	409.5	458.4	459.2	1.8
tadZ	Pilus assembly protein CpaE	334.8	295.2	389.4	339.8	731.3	599.8	620.3	650.5	1.9
rcpB	Pilus assembly protein FlpD	90.3	167.8	210.9	156.4	246.4	269.6	245.0	253.7	1.6
rcpA	Type II and III secretion system protein	285.2	287.0	239.1	270.4	520.7	475.6	474.8	490.4	1.8
rcpC	Flp pilus assembly protein CpaB	132.1	210.1	223.5	188.6	148.3	117.4	148.0	137.9	0.7
tadV	Pilus assembly protein PilD	488.9	578.9	442.1	503.3	1270.9	781.3	843.2	965.1	1.9
flp	Flp family type IVb pilin	110.1	208.4	224.1	180.9	514.6	394.3	341.9	416.9	2.3

Supplementary table 3.11 Expression data (normalized counts) for genes encoding a PTS system for N-Acetylgalactosamine utilization of strains SA5Y and SL5Y.

		2 h growth				6 h growth				Ratio
Gene	**Gene product**	**Repl 1**	**Repl 2**	**Repl 3**	**Average**	**Repl 1**	**Repl 2**	**Repl 3**	**Average**	**6h/ 2h**
Strain SA5Y										
gatY (agaY-like)	D-tagatose-1,6-bisphosphate aldolase subunit GatY	34.6	31.7	24.3	30.2	96.5	105.2	96.3	99.3	3.3
nagA	N-acetylglucosamine-6-phosphate deacetylase	17.3	17.1	17.0	17.1	82.8	74.5	70.9	76.1	4.4
manX	PTS system mannose-specific EIIAB component	23.0	14.7	30.8	22.8	52.8	55.7	57.1	55.2	2.4
manZ	PTS system mannose-specific EIID component	24.0	16.6	14.0	18.2	131.3	138.5	164.4	144.7	8.0
agaC	N-acetylgalactosamine permease IIC component	19.8	18.8	32.2	23.6	98.5	101.0	132.4	110.6	4.7
sorB	PTS system sorbose-specific EIIB component	6.2	2.4	2.3	3.6	22.9	37.5	25.1	28.5	7.8
agaS	D-galactosamine-6-phosphate deaminase	14.6	9.9	3.5	9.3	65.9	53.4	69.9	63.0	6.8
kbaZ	D-tagatose-1,6-bisphosphate aldolase subunit	20.0	6.2	9.1	11.8	41.5	49.7	39.7	43.6	3.7
srlR_1	Glucitol operon repressor	12.4	6.8	12.1	10.4	24.1	30.3	26.1	26.8	2.6
srlR_2	Glucitol operon repressor	297.7	175.5	221.7	231.6	829.5	998.2	832.6	886.8	3.8
Strain SL5Y										
gatY (agaY-like)	D-tagatose-1,6-bisphosphate aldolase subunit GatY	38.2	10.2	10.0	19.5	71.2	62.0	107.7	80.3	4.1
nagA	N-acetylglucosamine-6-phosphate deacetylase	19.1	9.0	13.0	13.7	48.7	60.8	48.0	52.5	3.8
manX	PTS system mannose-specific EIIAB component	21.6	5.3	8.5	11.8	32.4	38.8	45.1	38.8	3.3
manZ	PTS system mannose-specific EIID component	28.3	8.8	13.7	16.9	70.3	92.5	77.9	80.3	4.7
agaC	N-acetylgalactosamine permease IIC component	23.1	8.1	7.6	12.9	64.9	98.8	87.9	83.9	6.5
sorB	PTS system sorbose-specific EIIB component	6.3	1.1	2.4	3.3	20.5	15.5	16.4	17.5	5.3
agaS	D-galactosamine-6-phosphate deaminase	29.1	10.1	10.6	16.6	23.2	44.1	46.3	37.9	2.3
kbaZ	D-tagatose-1,6-bisphosphate aldolase subunit	16.0	7.5	10.5	11.3	27.2	25.5	28.8	27.2	2.4
srlR_1	Glucitol operon repressor	16.2	11.7	20.5	16.1	14.3	16.9	14.2	15.1	0.9
srlR_2	Glucitol operon repressor	575.0	552.4	447.4	524.9	888.0	979.6	805.2	890.9	1.7

Supplementary table 3.12 Expression data (normalized counts) for operons encoding piscibactin biosynthesis and for glycine betaine production. RNAseq data exemplified for strains L6G, SP7G, and SA10G.

		2 h growth				6 h growth				Ratio
Gene	**Gene product**	**Repl 1**	**Repl 2**	**Repl 3**	**Average**	**Repl 1**	**Repl 2**	**Repl 3**	**Average**	**6h/ 2h**
Strain L6G										
-	Helix-turn-helix transcriptional regulator	23030.9	21401.2	16519.6	20317.2	1836.0	2183.3	1926.7	1982.0	0.1
-	Helix-turn-helix transcriptional regulator	5928.4	5805.7	5091.8	5608.6	792.9	879.1	825.8	832.6	0.1
(***frpA***)	TonB-dependent receptor	36677.7	32761.9	32186.6	33875.4	1132.0	1518.0	1299.8	1316.6	0.0
(***irp8***)	MFS transporter	15962.5	13471.1	12932.0	14121.9	851.7	1000.7	843.8	898.7	0.1
(***irp2***)	Amino acid adenylation domain-containing protein	15159.0	13871.1	13896.7	14308.9	545.1	729.3	765.5	680.0	0.0
(***irp1***)	Amino acid adenylation domain-containing protein	10320.7	9579.1	8986.2	9628.7	1504.0	1332.6	1560.4	1465.7	0.2
(***irp3***)	Gfo/Idh/MocA family oxidoreductase	1039.1	903.0	869.5	937.2	237.0	171.1	213.5	207.2	0.2
(***irp4***)	Thioesterase	717.6	591.2	638.4	649.1	123.9	109.8	129.8	121.2	0.2
(***irp9***)	Salicylate synthase	2107.9	1738.7	1907.6	1918.1	613.6	445.0	581.8	546.8	0.3
(***irp5***)	AMP-binding protein	1838.1	1508.3	1475.4	1607.3	513.3	366.7	433.2	437.7	0.3
betT	BCCT family transporter	130.8	161.4	179.6	157.3	1386.7	3337.5	2342.8	2355.7	15.0
betI	Transcriptional regulator	48.3	48.3	95.1	63.9	397.4	738.8	565.2	567.1	8.9
betB	Betaine-aldehyde dehydrogenase	146.2	133.9	182.2	154.1	1599.4	2748.5	2174.3	2174.0	14.1
betA	Choline dehydrogenase	245.2	226.9	265.4	245.8	2595.3	4220.1	3716.1	3510.5	14.3
-	ABC transporter ATP-binding protein	15690.2	14048.5	11853.5	13864.0	2320.8	2159.9	2325.4	2268.7	0.2
-	ABC transporter ATP-binding protein	21131.4	19309.3	15628.9	18689.9	2449.0	2304.1	2375.1	2376.1	0.1
Strain SP7G										
-	Helix-turn-helix transcriptional regulator	12133.8	17478.7	12457.1	14023.2	1928.6	1064.5	1095.8	1362.9	0.1
-	Helix-turn-helix transcriptional regulator	3165.0	4068.6	2822.3	3351.9	889.6	617.4	718.8	742.0	0.2
(***frpA***)	TonB-dependent receptor	11472.1	21893.7	13099.5	15488.4	420.3	487.5	416.0	441.3	0.0
(***irp8***)	MFS transporter	5528.6	8538.2	5330.9	6465.9	294.8	310.8	236.7	280.8	0.0
(***irp2***)	Amino acid adenylation domain-containing protein	7357.5	7244.0	4487.0	6362.9	181.6	239.8	159.2	193.6	0.0
(***irp1***)	Amino acid adenylation domain-containing protein	6807.0	4321.9	2892.0	4673.6	669.2	941.8	644.6	751.9	0.2
(***irp3***)	Gfo/Idh/MocA family oxidoreductase	762.9	511.8	332.3	535.7	176.4	171.0	175.9	174.4	0.3
(***irp4***)	Thioesterase	628.0	355.7	205.6	396.4	69.6	127.0	57.5	84.7	0.2
(***irp9***)	Salicylate synthase	1712.2	1107.5	742.3	1187.3	288.3	430.6	315.1	344.7	0.3
(***irp5***)	AMP-binding protein	1330.2	1021.0	674.4	1008.6	202.0	296.0	217.3	238.4	0.2
betT	BCCT family transporter	6655.0*	170.8	264.4	217.6**	862.2	769.4	769.1	800.2	3.7
betI	Transcriptional regulator	1399.6*	60.4	71.7	66.1**	268.4	281.1	285.2	278.2	4.2

betB	Betaine-aldehyde dehydrogenase	3161.6*	130.2	173.2	151.7**	907.0	1003.4	908.0	939.5	6.2
betA	Choline dehydrogenase	2648.4*	166.8	216.7	191.7**	1864.5	1853.0	1740.0	1819.1	9.5
-	ABC transporter ATP-binding protein	4957.9	10204.3	6772.4	7311.5	885.7	808.0	807.2	833.6	0.1
-	ABC transporter ATP-binding protein	6861.8	13619.4	10185.3	10222.2	1310.1	1244.7	1174.4	1243.1	0.1
Strain SA10G										
-	Helix-turn-helix transcriptional regulator	20056.3	32279.9	21788.6	24708.3	2784.1	2293.5	2517.5	2531.7	0.1
-	Helix-turn-helix transcriptional regulator	3920.2	5914.1	4451.8	4762.1	900.9	1089.2	1112.6	1034.2	0.2
(***frpA***)	TonB-dependent receptor	22708.9	30632.2	34062.9	29134.7	2009.5	1941.9	2035.4	1995.6	0.1
(***irp8***)	MFS transporter	12003.8	11923.0	13098.0	12341.6	1220.6	1113.4	1146.3	1160.1	0.1
(***irp2***)	Amino acid adenylation domain-containing protein	9475.9	9491.2	13512.4	10826.5	1134.6	852.5	1109.2	1032.1	0.1
(***irp1***)	Amino acid adenylation domain-containing protein	8036.1	6081.8	9610.2	7909.4	2201.3	1305.3	1554.4	1687.0	0.2
(***irp3***)	Gfo/Idh/MocA family oxidoreductase	1024.7	787.2	1015.0	942.3	346.7	251.7	275.6	291.3	0.3
(***irp4***)	Thioesterase	545.2	570.8	813.0	643.0	180.4	129.9	137.0	149.1	0.2
(***irp9***)	Salicylate synthase	1029.8	1721.9	2405.0	1718.9	777.9	488.2	467.9	578.0	0.3
(***irp5***)	AMP-binding protein	1437.2	1360.4	1803.0	1533.6	593.2	380.9	389.0	454.3	0.3
betT	BCCT family transporter	399.4	193.7	131.5	241.6	3072.8	3815.3	4043.9	3644.0	15.1
betI	Transcriptional regulator	63.1	60.1	52.5	58.6	653.9	889.5	627.8	723.8	12.4
betB	Betaine-aldehyde dehydrogenase	227.1	121.7	187.7	178.8	2590.3	2347.8	2498.7	2478.9	13.9
betA	Choline dehydrogenase	246.8	246.3	259.2	250.8	4202.1	3569.7	3650.8	3807.5	15.2
-	ABC transporter ATP-binding protein	11476.9	13702.7	10985.1	12054.9	2703.7	1998.2	2105.3	2269.1	0.2
-	ABC transporter ATP-binding protein	16081.6	19883.5	16238.3	17401.2	3142.8	2521.5	2997.1	2887.2	0.2
-	Helix-turn-helix transcriptional regulator	20056.3	32279.9	21788.6	24708.3	2784.1	2293.5	2517.5	2531.7	0.1

*normalized counts of these genes in this replicate deviated considerably (in this case, more than 10-fold) from the other two replicates, and were therefore not included in the average (**) calculation.

Gene names in brackets according to nomenclature of the piscibactin cluster in the original strain where it was first described (*P. damselae* subsp. *piscicida* DI21, plasmid pPHDP70).

Supplementary table 3.13 Expression data (normalized counts) for *lux* operon. RNAseq data exemplified for strains L6G and SA10G.

		2 h growth				6 h growth				Ratio
Gene	**Gene product**	**Repl 1**	**Repl 2**	**Repl 3**	**Average**	**Repl 1**	**Repl 2**	**Repl 3**	**Average**	**6h/ 2h**
Strain L6G										
***fre* (*luxG*)**	NAD(P)H-flavin reductase	142.0	114.9	169.1	142.0	221.5	736.6	455.8	471.3	3.3
luxE	Long-chain fatty acid--CoA ligase	142.1	155.0	171.8	156.3	162.8	511.9	358.5	344.4	2.2
luxB	LLM class flavin-dependent oxidoreductase	226.1	301.9	353.0	293.7	1094.9	3700.9	2439.0	2411.6	8.2
luxA	LLM class flavin-dependent oxidoreductase	217.7	274.5	374.7	289.0	1041.0	2984.1	1892.6	1972.6	6.8
luxD	Acyl transferase	150.1	173.8	269.4	197.8	452.8	1566.1	984.0	1000.9	5.1
luxC	Aldehyde dehydrogenase family protein	445.1	470.7	824.8	580.2	482.2	1485.4	920.5	962.7	1.7
Strain SA10G										
***fre* (*luxG*)**	NAD(P)H-flavin reductase	128.4	166.6	195.2	163.4	159.1	261.5	218.5	213.0	1.3
luxE	Long-chain fatty acid--CoA ligase	231.0	368.2	408.0	335.7	211.3	357.5	377.9	315.6	0.9
luxB	LLM class flavin-dependent oxidoreductase	1123.3	1856.6	1440.7	1473.5	2370.1	3282.7	3122.5	2925.1	2.0
luxA	LLM class flavin-dependent oxidoreductase	988.5	1145.7	1283.6	1139.3	1702.2	2138.3	2743.5	2194.7	1.9
luxD	Acyl transferase	809.5	780.1	966.4	852.0	1166.9	1393.1	1470.4	1343.5	1.6
luxC	Aldehyde dehydrogenase family protein	1187.8	1475.0	1767.2	1476.7	510.2	962.8	798.8	757.3	0.5

Supplementary table 3.14 Expression data (normalized counts) for *nan-nag* and T3SS operons. RNAseq data exemplified for strains SA5Y and SL5Y.

		2 h growth				6 h growth				Ratio
Gene	**Gene product#**	**Repl 1**	**Repl 2**	**Repl 3**	**Average**	**Repl 1**	**Repl 2**	**Repl 3**	**Average**	**6h/ 2h**
Strain SA5Y										
nanH	Sialidase	5853.1	1354.9	1929.0	3045.7	1362.6	1272.3	1114.7	1249.9	0.4
nagA	N-acetylglucosamine-6-phosphate deacetylase	2352.2	639.0	805.2	1265.5	861.9	804.7	770.7	812.4	0.6
nanK	N-acetylmannosamine kinase	2946.9	921.8	1422.2	1763.7	215.0	210.6	167.0	197.5	0.1
nanE	Putative N-acetylmannosamine-6-phosphate 2-epimerase	3144.3	1695.7	2133.2	2324.4	110.5	104.9	68.6	94.7	0.0
siaP	Sialic acid-binding periplasmic protein	6559.3	1577.5	2451.2	3529.3	262.7	255.6	240.1	252.8	0.1
siaQ	Sialic acid TRAP transporter small permease protein	2702.4	359.5	477.5	1179.8	99.5	89.9	62.9	84.1	0.1
siaM	Sialic acid TRAP transporter large permease protein	4593.2	749.3	855.2	2065.9	334.2	328.7	376.1	346.3	0.2
nanA	N-acetylneuraminate lyase	2897.3	759.3	950.6	1535.8	344.8	369.0	305.1	339.6	0.2
ybbH	Putative HTH-type transcriptional regulator	4200.7	2355.3	2279.9	2945.3	1976.3	2048.6	2021.5	2015.5	0.7
nanM	N-acetylneuraminate epimerase	3277.5	1478.3	1756.1	2170.6	929.2	874.9	792.6	865.5	0.4
nanM	N-acetylneuraminate epimerase	1768.2	890.7	1016.8	1225.3	656.6	822.9	650.6	710.0	0.6
-	Hypothetical	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0
-	Hypothetical	953.5	1100.8	1019.0	1024.4	261.6	281.0	221.3	254.6	0.2
-	Hypothetical	192.7	205.4	217.6	205.2	11.9	9.6	18.3	13.3	0.1
-	Hypothetical	58.4	74.7	75.7	69.6	4.9	2.5	0.0	2.5	0.0
-	Hypothetical	30.0	45.7	37.6	37.8	3.0	2.1	6.9	4.0	0.1
-	Hypothetical	12.4	10.6	25.1	16.0	2.0	2.1	5.7	3.3	0.2
-	Hypothetical	5.3	0.8	1.1	2.4	0.0	0.0	0.0	0.0	0.0
-	Arylesterase	1276.1	1522.0	1482.5	1426.9	92.6	138.1	120.0	116.9	0.1
tdh2	Thermostable direct hemolysin 2	127.3	106.0	117.4	116.9	10.0	19.3	6.9	12.0	0.1
-	SslE/AcfD family lipoprotein zinc metalloprotease	1815.8	1896.6	1825.1	1845.8	1189.0	1372.9	1174.2	1245.4	0.7
vopY	Predicted effector	433.9	438.6	452.4	441.6	212.0	357.6	282.3	284.0	0.6
vopK	Predicted effector	356.2	343.8	360.0	353.3	52.8	49.0	32.7	44.8	0.1
-	Hypothetical	66.3	75.0	62.7	68.0	8.0	9.6	11.4	9.7	0.1
vopG	Predicted effector	8.0	18.8	22.8	16.5	6.0	4.3	4.6	4.9	0.3
-	Hypothetical	11.5	13.9	23.9	16.4	6.0	2.1	0.0	2.7	0.2
vopF	Predicted effector	538.8	427.2	450.4	472.1	414.6	459.2	422.5	432.1	0.9
-	Hypothetical	344.7	328.5	371.5	348.2	1050.1	1121.0	908.7	1026.6	2.9
yscI	EscI/YscI/HrpB family type III secretion system inner rod protein	26.5	31.8	41.0	33.1	75.6	66.4	52.6	64.9	2.0
yscJ	Type III secretion apparatus lipoprotein YscJ/ HrcJ family	92.8	99.5	94.6	95.6	140.4	161.7	155.4	152.5	1.6
-	Hypothetical	94.6	99.5	137.9	110.6	51.8	71.7	65.1	62.9	0.6
-	Hypothetical (TRP repeats; chaperone?)	58.3	56.3	66.1	60.2	32.8	28.9	10.3	24.0	0.4
-	(Chaperone?)	482.2	393.7	477.9	451.3	158.5	237.7	154.7	183.7	0.4

-	(Chaperone?)	531.2	508.0	497.0	512.1	133.6	175.3	109.4	139.4	0.3
vopB 2	Type III secretion system translocator protein	758.7	732.0	818.3	769.7	140.2	148.9	139.6	142.9	0.2
-	Type III secretion system protein CDS	895.4	796.4	991.7	894.5	345.2	375.6	354.2	358.4	0.4
-	ATP dependent exonuclease	214.7	169.6	216.5	200.3	25.9	56.7	70.9	51.2	0.3
vopI	Predicted effector	58.7	50.4	86.9	65.3	4.2	13.6	12.1	10.0	0.2
-	Dimethyladenosine transferase	173.2	152.4	192.6	172.7	40.8	47.1	43.4	43.8	0.3
-	Hypothetical	180.3	148.4	200.6	176.4	38.8	43.9	32.0	38.2	0.2
vopM	Predicted effector	1566.3	1417.9	1464.4	1482.9	2386.6	2196.0	1904.5	2162.4	1.5
-	Hypothetical	198.8	178.5	202.8	193.4	82.6	80.3	60.6	74.5	0.4
ssaV	Secretion system apparatus protein	734.6	695.9	676.4	702.3	854.9	796.7	787.4	813.0	1.2
yscU	Yop proteins translocation protein U	206.8	239.2	259.2	235.1	159.1	148.7	103.8	137.2	0.6
vopA	OmpA family protein	244.1	218.8	215.1	226.0	176.6	194.7	160.7	177.3	0.8
VPA1 352	Chaperone VPA1352	173.5	181.6	190.2	181.8	98.2	126.4	105.4	110.0	0.6
vopH	Predicted effector	580.2	535.6	565.1	560.3	429.2	500.4	381.5	437.1	0.8
fliM	FliM	508.1	441.0	527.6	492.3	1311.9	1524.6	1312.1	1382.9	2.8
fliN	FliM/FliN family flagellar motor switch protein	731.7	658.7	735.0	708.5	914.8	1129.6	906.4	983.6	1.4
toxR/ vttrB	Transcriptional activator ToxR/VttrB	282.8	310.6	302.0	298.5	211.0	173.4	190.9	191.8	0.6
-	Hypothetical	620.4	589.4	696.2	635.3	913.8	1123.1	1042.4	1026.4	1.6
yscR	EscR/YscR/HrcR family type III secretion system export apparatus protein	99.9	133.7	156.1	129.9	44.8	60.0	58.3	54.3	0.4
vscT2	type III secretion system apparatus protein VscT2	80.4	102.7	121.9	101.7	63.7	74.9	50.3	63.0	0.6
VPA1 340	VPA1340 family putative T3SS effector	54.8	53.0	78.6	62.1	60.7	68.5	35.4	54.9	0.9
spiA	Type III secretion system outer membrane protein	510.7	501.6	548.9	520.4	728.8	722.4	623.9	691.7	1.3
yscN	putative ATP synthase YscN	568.0	584.6	563.7	572.1	793.2	752.0	661.7	735.6	1.3
VPA1 337	VPA1337 family putative T3SS effector	272.6	225.8	258.4	252.2	77.9	75.5	71.8	75.1	0.3
-	Hypothetical	505.5	522.6	551.5	526.5	79.6	123.1	92.6	98.4	0.2
fliQ	Flagellar biosynthetic protein	124.6	103.5	144.7	124.3	14.9	15.0	20.6	16.8	0.1
-	Hypothetical	117.3	154.1	138.0	136.5	15.4	24.3	21.3	20.4	0.1
-	Hypothetical	3051.5	2816.6	3124.6	2997.6	405.1	527.8	387.5	440.1	0.1
hilA/ vttrA	Transcriptional regulator HilA/ VttrA	6220.5	6134.6	5985.9	6113.7	1091.9	1550.3	1421.8	1354.7	0.2
vopX	Predicted effector	689.3	790.8	773.7	751.3	431.0	578.2	489.2	499.5	0.7
-	Hypothetical	152.9	199.7	186.9	179.8	123.4	114.6	120.0	119.3	0.7
yopE	Outer membrane virulence protein YopE/ VopE	74.2	112.5	121.9	102.9	5.0	9.6	13.7	9.4	0.1
acfA	AcfA family outer membrane beta-barrel protein	125.5	144.3	170.9	146.9	17.9	22.5	22.9	21.1	0.1
intA_ 1	Prophage integrase	1780.6	1232.1	1628.3	1547.0	3945.8	4071.7	4211.1	4076.2	2.6

#Gene annotations and corresponding products were obtained by a combination of automated annotation pipelines (PROKKA and PGAP), BLASTp searches and genome comparison with the island found in *V. cholerae* AM-19226 (contigs AATY02000003 and AATY02000004).

Gene	Gene product#	2 h growth				6 h growth				Ratio
		Repl 1	Repl 2	Repl 3	Average	Repl 1	Repl 2	Repl 3	Average	6h/ 2h
Strain SL5Y										
nanH	Sialidase	20511*	3605.9	3061.1	3333.5**	1335.9	1277.7	1145.5	1253.0	0.1
nagA	N-acetylglucosamine-6-phosphate deacetylase	5763.3*	1258.8	769.3	1014.0**	719.1	707.2	706.7	711.0	0.3
nanK	N-acetylmannosamine kinase	5694.7*	1474.9	1437.3	1456.1**	203.2	206.4	178.8	196.1	0.1
nanE	Putative N-acetylmannosamine-6-phosphate 2-epimerase	5303.9*	2101.1	2635.8	2368.4**	133.5	102.2	89.3	108.4	0.0
siaP	Sialic acid-binding periplasmic protein	15525*	3866.3	3164.4	3515.3**	256.3	257.5	271.7	261.8	0.0
siaQ	Sialic acid TRAP transporter small permease protein	4374.8*	1318.2	554.6	936.4**	84.3	66.6	72.8	74.6	0.0
siaM	Sialic acid TRAP transporter large permease protein	9756.5*	2276.1	1018.8	1647.4**	275.8	333.2	334.7	314.6	0.1
nanA	N-acetylneuraminate lyase	7532.7*	1277.1	855.6	1066.3**	378.1	315.2	294.1	329.1	0.1
ybbH	Putative HTH-type transcriptional regulator	6099.7*	2872.3	1918.5	2395.4**	1703.9	1699.3	1656.6	1686.6	0.5
nanM	N-acetylneuraminate epimerase	6406.6*	2106.1	1466.0	1786.0**	678.6	691.6	604.0	658.1	0.2
nanM	N-acetylneuraminate epimerase	3174.1*	1134.6	728.0	931.3**	520.3	545.3	525.8	530.5	0.3
-	Hypothetical	0.0	2.1	0.0	0.7	0.0	0.0	0.0	0.0	0.0
-	Hypothetical	1144.5	746.3	602.7	831.2	232.2	250.5	230.3	237.7	0.3
-	Hypothetical	94.0	254.8	129.6	159.4	14.1	18.9	16.4	16.4	0.1
-	Hypothetical	27.0	64.6	39.5	43.7	4.0	1.2	6.1	3.8	0.1
-	Hypothetical	26.7	38.1	21.8	28.8	3.2	1.1	7.2	3.8	0.1
-	Hypothetical	12.7	13.7	8.5	11.6	1.1	0.0	1.0	0.7	0.1
-	Arylesterase	727.7	1307.6	1056.0	1030.4	107.0	99.8	98.5	101.8	0.1
tdh2	Thermostable direct hemolysin 2	73.7	171.2	64.2	103.0	8.6	6.7	9.2	8.2	0.1
-	SslE/AcfD family lipoprotein zinc metalloprotease	1656.4	1759.6	1262.6	1559.5	1112.0	1133.8	995.4	1080.4	0.7
vopY	Predicted effector	330.8	456.7	278.5	355.3	278.9	307.3	282.1	289.4	0.8
vopK	Predicted effector	264.0	496.5	299.4	353.3	37.0	46.9	42.4	42.1	0.1
-	Hypothetical	38.7	82.5	47.2	56.1	13.0	3.3	7.2	7.8	0.1
vopG	Predicted effector	10.2	16.9	8.5	11.8	5.4	4.4	2.1	4.0	0.3
-	Hypothetical	8.9	7.4	4.8	7.0	4.3	0.0	5.1	3.2	0.4
vopF	Predicted effector	445.3	546.1	414.1	468.5	469.1	453.6	409.7	444.1	0.9
-	Hypothetical	353.7	526.4	350.0	410.0	1080.0	1210.3	1070.9	1120.4	2.7
yscI	EscI/YscI/HrpB family type III secretion system inner rod protein	30.5	69.8	55.7	52.0	61.6	61.0	55.4	59.3	1.1
yscJ	Type III secretion apparatus lipoprotein YscJ/ HrcJ family	126.4	209.3	141.7	159.1	233.5	309.5	224.7	255.9	1.6
-	Hypothetical	115.6	206.1	135.6	152.4	137.3	103.2	115.9	118.8	0.8
-	Hypothetical (TRP repeats; chaperone?)	123.2	212.5	139.3	158.3	81.1	109.8	88.2	93.0	0.6
-	(Chaperone?)	520.0	616.0	408.5	514.8	542.0	596.0	468.5	535.5	1.0
vopB 2	Type III secretion system translocator protein	475.2	709.6	523.4	569.4	197.7	167.3	139.3	168.1	0.3
-	Type III secretion system protein CDS	629.5	955.7	714.2	766.5	362.8	386.9	386.0	378.5	0.5
-	ATP dependent exonuclease	118.1	189.2	106.6	138.0	46.5	42.2	49.2	46.0	0.3

(***vopI***)	Predicted effector	30.7	34.0	56.4	40.4	25.6	17.1	16.9	19.9	0.5
-	Dimethyladenosine transferase	81.3	172.3	100.5	118.0	37.8	46.6	48.2	44.2	0.4
vopM	Predicted effector	1377.9	1267.2	1291.4	1312.2	1704.9	1578.7	1572.8	1618.8	1.2
-	Hypothetical	104.1	125.8	135.6	121.9	68.1	74.3	84.1	75.5	0.6
ssaV	Secretion system apparatus protein	559.5	756.1	451.1	588.9	689.4	664.0	585.9	646.5	1.1
yscU	Yop proteins translocation protein U	142.5	273.2	109.8	175.1	111.3	130.8	100.9	114.3	0.7
vopA	OmpA family protein	185.9	197.2	155.0	179.4	143.1	181.5	142.1	155.6	0.9
VPA1 352	Chaperone VPA1352	125.8	171.9	127.6	141.8	118.9	114.2	103.3	112.1	0.8
vopH	Predicted effector	428.0	644.8	413.4	495.4	385.0	386.7	312.3	361.3	0.7
fliM	FliM	449.5	550.7	434.8	478.3	1377.3	1391.1	1209.4	1326.0	2.8
fliN	FliM/FliN family flagellar motor switch protein	582.2	676.5	508.6	589.1	795.7	880.8	930.4	869.0	1.5
vttrB	Transcriptional activator ToxR	259.7	287.5	228.9	258.7	176.2	170.8	151.8	166.3	0.6
-	Hypothetical	519.6	551.4	542.0	537.7	785.3	896.6	1032.2	904.7	1.7
yscR	EscR/YscR/HrcR family type III secretion system export apparatus protein	77.5	127.9	88.4	97.9	37.8	58.8	47.2	47.9	0.5
vscT2	type III secretion system apparatus protein VscT2	78.7	107.8	75.1	87.2	56.2	46.6	51.3	51.4	0.6
VPA1 340	VPA1340 family putative T3SS effector	27.9	65.5	41.2	44.9	44.3	56.6	47.2	49.4	1.1
spiA	Type III secretion system outer membrane protein	518.8	603.8	424.1	515.6	619.8	635.1	540.8	598.6	1.2
yscN	putative ATP synthase YscN	554.3	678.9	485.9	573.1	647.6	608.6	585.8	614.0	1.1
VPA13 37	VPA1337 family putative T3SS effector	143.5	180.8	123.5	149.3	27.0	34.4	24.6	28.7	0.2
-	Hypothetical	358.1	525.4	356.0	413.2	74.6	81.0	91.3	82.3	0.2
fliQ	Flagellar biosynthetic protein	71.7	170.2	53.3	98.4	7.6	12.2	10.3	10.0	0.1
-	Hypothetical	87.0	132.1	73.9	97.7	25.9	12.2	18.5	18.9	0.2
-	Hypothetical	1397.5	4120.5	1823.8	2447.3	286.5	360.5	330.3	325.8	0.1
hilA (vttrA)	Transcriptional regulator HilA/ VttrA	2582.3	6659.6	3691.2	4311.1	1136.3	1152.6	1019.6	1102.8	0.3
vopX	Predicted effector	461.6	944.0	554.6	653.4	396.8	442.6	407.2	415.5	0.6
-	Hypothetical	94.6	274.8	128.4	165.9	97.3	112.0	100.5	103.3	0.6
yopE	Outer membrane virulence protein YopE/ VopE	56.5	108.9	66.6	77.3	6.5	5.5	3.1	5.0	0.1
acfA	AcfA family outer membrane beta-barrel protein	81.3	226.2	104.1	137.2	16.2	18.9	19.5	18.2	0.1
-	Transposase	97.1	203.0	116.3	138.8	51.9	51.0	53.3	52.1	0.4
intA_ 1	Prophage integrase	1237.2	1553.1	1182.5	1324.3	862.9	861.9	807.1	844.0	0.6

#Gene annotations and corresponding products were obtained by a combination of automated annotation pipelines (PROKKA and PGAP), BLASTp searches and genome comparison with the island found in *V. cholerae* AM-19226 (contigs AATY02000003 and AATY02000004).

*normalized counts of these genes in this replicate deviated considerably (in this case, around 5-fold or more) from the other two replicates, and were therefore not included in the average (**) calculation.

3.8 REFERENCES

1. Gevers,D., Cohan,F.M., Lawrence,J.G., Spratt,B.G., Coenye,T., Feil,E.J., Stackebrandt,E., Peer,Y. Van De, Vandamme,P., Thompson,F.L., *et al.* (2005) Re-evaluating prokaryotic species. *Nat. Rev. Microbiol.*, **3**, 733–739.

2. Shapiro,B.J. and Polz,M.F. (2014) Ordering microbial diversity into ecologically and genetically cohesive units. *Trends Microbiol.*, **22**, 235–247.

3. Ford Doolittle,W. and Papke,R.T. (2006) Genomics and the bacterial species problem. *Genome Biol.*, **7**.

4. Bobay,L.M. (2020) The prokaryotic species concept and challenges. In Tettelin,H., Medini,D. (eds), *The Pangenome: Diversity, Dynamics and Evolution of Genomes*. Springer International Publishing, pp. 21–49.

5. Frost,L.S., Leplae,R., Summers,A.O. and Toussaint,A. (2005) Mobile genetic elements: The agents of open source evolution. *Nat. Rev. Microbiol.*, **3**, 722–732.

6. Toussaint,A. and Chandler,M. (2012) Prokaryote genome fluidity: toward a system approach of the mobilome. In Helden,J. van, Toussaint,A., Thieffry,D. (eds), *Bacterial Molecular Networks: Methods and Protocols*. Springer International Publishing, pp. 57–80.

7. Innamorati,K.A., Earl,J.P., Aggarwal,S.D., Ehrlich,G.D. and Hiller,N.L. (2020) The bacterial guide to designing a diversified gene portfolio. In Tettelin,H., Medini,D. (eds), *The Pangenome: Diversity, Dynamics and Evolution of Genomes*. Springer International Publishing, pp. 51–87.

8. Syvanen,M. (2012) Evolutionary implications of horizontal gene transfer. *Annu. Rev. Genet.*, **46**, 341–358.

9. Dobrindt,U., Hochhut,B., Hentschel,U. and Hacker,J. (2004) Genomic islands in pathogenic and environmental microorganisms. *Nat. Rev. Microbiol.*, **2**, 414–424.

10. Toussaint,A. and Merlin,C. (2002) Mobile elements as a combination of functional modules. *Plasmid*, **47**, 26–35.

11. Osborn,A.M. and Böltner,D. (2002) When phage, plasmids, and transposons collide: Genomic islands, and conjugative- and mobilizable-transposons as a mosaic continuum. *Plasmid*, **48**, 202–212.

12. Ochman,H., Lawrence,J.G. and Grolsman,E.A. (2000) Lateral gene transfer and the nature of bacterial innovation. *Nature*, **405**, 299–304.

13. Olendzeski,L. and Gogarten,J.P. (2009) Gene transfer: who benefits? In Gogarten,M.B., Gogarten,J.P., Olendzeski,L. (eds), *Horizontal gene transfer: genomes in flux*. Springer International Publishing, pp. 3–9.

14. Schmidt,H. and Hensel,M. (2004) Pathogenicity islands in bacterial pathogenesis. *Clin. Microbiol. Rev.*, **17**, 14–56.

15. Ahmed,N., Dobrindt,U., Hacker,J. and Hasnain,S.E. (2008) Genomic fluidity and pathogenic bacteria: applications in diagnostics, epidemiology and intervention. *Nat. Rev. Micro*, **6**, 387–394.

16. Faruque,S.M., Albert,M.J. and Mekalanos,J.J. (1998) Epidemiology, genetics, and ecology of toxigenic *Vibrio cholerae*. *Microbiol. Mol. Biol. Rev.*, **62**, 1301–14.

17. Cottingham,K.L., Chiavelli,D.A. and Taylor,R.K. (2003) Environmental microbe and human pathogen: the ecology and microbiology of *Vibrio cholerae*. *Front Ecol Environm*, **1**, 80–86.

18. World Health Organization (WHO) (2006) Communicable diseases following natural disasters Risk assessment and priority interventions.

19. Clemens,J.D., Nair,G.B., Ahmed,T., Qadri,F. and Holmgren,J. (2017) Cholera. *Lancet*, **390**, 1539–1549.

20. Robins,W.P. and Mekalanos,J.J. (2014) Genomic science in understanding cholera outbreaks and evolution of *Vibrio cholerae* as a human pathogen. In Nair,G., Takeda,Y. (eds), *Cholera Outbreaks.*

Current Topics in Microbiology and Immunology, vol 379. Springer, Berlin, Heidelberg, pp. 211–229.

21. Waldor,M.K. and Mekalanos,J.J. (1996) Lysogenic conversion by a filamentous phage encoding cholera toxin. *Science*, **272**, 1910–1914.

22. Taylor,R.K., Miller,V.L., Furlong,D.B. and Mekalanos,J.J. (1987) Use of *phoA* gene fusions to identify a pilus colonization factor coordinately regulated with cholera toxin. *Proc. Natl. Acad. Sci.*, **84**, 2833–2837.

23. Karaolis,D.K.R., Johnson,J.A., Bailey,C.C., Boedeker,E.C., Kaper,J.B. and Reeves,P.R. (1998) A *Vibrio cholerae* pathogenicity island associated with epidemic and pandemic strains. *Proc. Natl. Acad. Sci. U. S. A.*, **95**, 3134–3139.

24. Jermyn,W.S. and Boyd,E.F. (2002) Characterization of a novel *Vibrio* pathogenicity island (VPI-2) encoding neuraminidase (*nanH*) among toxigenic *Vibrio cholerae* isolates. *Microbiology*, **148**, 3681–3693.

25. Murphy,R.A. and Boyd,E.F. (2008) Three pathogenicity islands of *Vibrio cholerae* can excise from the chromosome and form circular intermediates. *J. Bacteriol.*, **190**, 636–647.

26. Boyd,E.F., Carpenter,M.R., Chowdhury,N., Cohen,A.L., Haines-Menges,B.L., Kalburge,S.S., Kingston,J.J., Lubin,J.B., Ongagna-Yhombi,S.Y. and Whitaker,W.B. (2015) Post-genomic analysis of members of the family Vibrionaceae. *Microbiol. Spectr.*, **3**, VE-0009-2014.

27. Kim,E.J., Lee,C.H., Nair,G.B. and Kim,D.W. (2015) Whole-genome sequence comparisons reveal the evolution of *Vibrio cholerae* O1. *Trends Microbiol.*, **23**, 479–489.

28. Dziejman,M., Balon,E., Boyd,D., Fraser,C.M., Heidelberg,J.F. and Mekalanos,J.J. (2002) Comparative genomic analysis of *Vibrio cholerae*: Genes that correlate with cholera endemic and pandemic disease. *Proc. Natl. Acad. Sci.*, **99**, 1556–1561.

29. O'Shea,Y.A., Finnan,S., Reen,F.J., Morrissey,J.P., O'Gara,F. and Boyd,E.F. (2004) The *Vibrio* seventh pandemic island-II is a 26.9 kb genomic island present in *Vibrio cholerae* El Tor and O139 serogroup isolates that shows homology to a 43.4 kb genomic island in *V. vulnificus*. *Microbiology*, **150**, 4053–4063.

30. Davies,B.W., Bogard,R.W., Young,T.S. and Mekalanos,J.J. (2012) Coordinated regulation of accessory genetic elements produces cyclic di-nucleotides for *V. cholerae* virulence. *Cell*, **149**, 358–370.

31. Cohen,D., Melamed,S., Millman,A., Shulman,G., Oppenheimer-Shaanan,Y., Kacen,A., Doron,S., Amitai,G. and Sorek,R. (2019) Cyclic GMP–AMP signalling protects bacteria against viral infection. *Nature*, **574**, 691–695.

32. Le Roux,F. and Blokesch,M. (2018) Eco-evolutionary dynamics linked to horizontal gene transfer in Vibrios. *Annu. Rev. Microbiol.*, **72**, 89–110.

33. Mutreja,A., Kim,D.W., Thomson,N.R., Connor,T.R., Lee,J.H., Kariuki,S., Croucher,N.J., Choi,S.Y., Harris,S.R., Lebens,M., *et al.* (2011) Evidence for several waves of global transmission in the seventh cholera pandemic. *Nature*, **477**, 462–465.

34. LeGault,K.N., Hays,S.G., Angermeyer,A., McKitterick,A.C., Johura,F., Sultana,M., Alam,M. and Seed,K.D. (2020) Temporal shifts in antibiotic resistance elements govern virus-pathogen conflicts. *bioRxiv*, https://doi.org/10.1101/2020.12.16.423150.

35. Boyd,E.F. and Waldor,M.K. (2002) Evolutionary and functional analyses of variants of the toxin-coregulated pilus protein TcpA from toxigenic *Vibrio cholerae* non-O1/nonO139 serogroup isolates. *Microbiology*, **148**, 1655–1666.

36. Rivera,I.N.G., Chun,J., Huq,A., Sack,R.B. and Colwell,R.R. (2001) Genotypes associated with virulence in environmental isolates of *Vibrio cholerae*. *Appl. Environ. Microbiol.*, **67**, 2421–2429.

37. Faruque,S.M., Kamruzzaman,M., Meraj,I.M., Chowdhury,N., Nair,G.B., Sack,R.B., Colwell,R.R. and

Sack,D.A. (2003) Pathogenic potential of environmental *Vibrio cholerae* strains carrying genetic variants of the toxin-coregulated pilus pathogenicity island. *Infect. Immun.*, **71**, 1020–1025.

38. Faruque,S.M., Chowdhury,N., Kamruzzaman,M., Dziejman,M., Rahman,M.H., Sack,D.A., Nair,G.B. and Mekalanos,J.J. (2004) Genetic diversity and virulence potential of environmental *Vibrio cholerae* population in a cholera-endemic area. *Proc. Natl. Acad. Sci.*, **101**, 2123–2128.

39. Gennari,M., Ghidini,V., Caburlotto,G. and Lleo,M.M. (2012) Virulence genes and pathogenicity islands in environmental *Vibrio* strains nonpathogenic to humans. *FEMS Microbiol. Ecol.*, **82**, 563–573.

40. Bernardy,E.E., Turnsek,M.A., Wilson,S.K., Tarr,C.L. and Hammer,B.K. (2016) Diversity of clinical and environmental isolates of *Vibrio cholerae* in natural tansformation and contact-dependent bacterial killing indicative of type VI secretion system activity. *Appl. Environ. Microbiol.*, **82**, 2833–2842.

41. Labbate,M., Orata,F.D., Petty,N.K., Jayatilleke,N.D., King,W.L., Kirchberger,P.C., Allen,C., Mann,G., Mutreja,A., Thomson,N.R., *et al.* (2016) A genomic island in *Vibrio cholerae* with VPI-1 site-specific recombination characteristics contains CRISPR-Cas and type VI secretion modules. *Sci. Rep.*, **6**, 36891.

42. Rapa,R.A., Islam,A., Monahan,L.G., Mutreja,A., Thomson,N., Charles,I.G., Stokes,H.W. and Labbate,M. (2015) A genomic island integrated into *recA* of *Vibrio cholerae* contains a divergent *recA* and provides multi-pathway protection from DNA damage. *Environ. Microbiol.*, **17**, 1090–1102.

43. Drebes Dörr,N.C. and Blokesch,M. (2020) Interbacterial competition and anti-predatory behavior of environmental *Vibrio cholerae* strains. *Environ. Microbiol.*, **22**, 4485–4504.

44. Keymer,D.P., Miller,M.C., Schoolnik,G.K. and Boehm,A.B. (2007) Genomic and phenotypic diversity of coastal *Vibrio cholerae* strains is linked to environmental factors. *Appl. Environ. Microbiol.*, **73**, 3705–3714.

45. Miller,M.C., Keymer,D.P., Avelar,A., Boehm,A.B. and Schoolnik,G.K. (2007) Detection and transformation of genome segments that differ within a coastal population of *Vibrio cholerae* strains. *Appl. Environ. Microbiol.*, **73**, 3695–3704.

46. Heidelberg,J.F., Elsen,J.A., Nelson,W.C., Clayton,R.A., Gwinn,M.L., Dodson,R.J., Haft,D.H., Hickey,E.K., Peterson,J.D., Umayam,L., *et al.* (2000) DNA sequence of both chromosomes of the cholera pathogen *Vibrio cholerae*. *Nature*, **406**, 477–483.

47. Matthey,N., Drebes Dörr,N.C. and Blokesch,M. (2018) Long-read-based genome sequences of pandemic and environmental *Vibrio cholerae* strains. *Microbiol. Resour. Announc.*, **7**.

48. Roux,S., Hallam,S.J., Woyke,T. and Sullivan,M.B. (2015) Viral dark matter and virus–host interactions resolved from publicly available microbial genomes. *Elife*, **4**, e08490.

49. Hampton,H.G., Watson,B.N.J. and Fineran,P.C. (2020) The arms race between bacteria and their phage foes. *Nature*, **577**, 327–336.

50. Bernheim,A. and Sorek,R. (2020) The pan-immune system of bacteria: antiviral defence as a community resource. *Nat. Rev. Microbiol.*, **18**, 113–119.

51. van Houte,S., Buckling,A. and Westra,E.R. (2016) Evolutionary ecology of prokaryotic immune mechanisms. *Microbiol. Mol. Biol. Rev.*, **80**, 745–763.

52. Zhou,X., He,X., Liang,J., Li,A., Xu,T., Kieser,T., Helmann,J.D. and Deng,Z. (2005) A novel DNA modification by sulphur. *Mol. Microbiol.*, **57**, 1428–1438.

53. Xu,T., Yao,F., Zhou,X., Deng,Z. and You,D. (2010) A novel host-specific restriction system associated with DNA backbone S-modification in *Salmonella*. *Nucleic Acids Res.*, **38**, 7133–7141.

54. Wang,L., Chen,S., Xu,T., Taghizadeh,K., Wishnok,J.S., Zhou,X., You,D., Deng,Z. and Dedon,P.C. (2007) Phosphorothioation of DNA in bacteria by *dnd* genes. *Nat. Chem. Biol.*, **3**, 709–710.

55. Wang,L., Jiang,S., Deng,Z., Dedon,P.C. and Chen,S. (2019) DNA phosphorothioate modification - A new

multi-functional epigenetic system in bacteria. *FEMS Microbiol. Rev.*, **43**, 109–122.

56. An,X., Xiong,W., Yang,Y., Li,F., Zhou,X., Wang,Z., Deng,Z. and Liang,J. (2012) A novel target of IscS in *Escherichia coli*: participating in DNA phosphorothioation. *PLoS One*, **7**, e51265.

57. Barrangou,R. and Horvath,P. (2012) CRISPR: New horizons in phage resistance and strain identification. *Annu. Rev. Food Sci. Technol.*, **3**, 143–162.

58. Amitai,G. and Sorek,R. (2016) CRISPR-Cas adaptation: Insights into the mechanism of action. *Nat. Rev. Microbiol.*, **14**, 67–76.

59. Koonin,E. V, Makarova,K.S. and Zhang,F. (2017) Diversity, classification and evolution of CRISPR-Cas systems. *Curr. Opin. Microbiol.*, **37**, 67–78.

60. Deveau,H., Garneau,J.E. and Moineau,S. (2010) CRISPR/Cas system and its role in phage-bacteria interactions. *Annu. Rev. Microbiol.*, **64**, 475–493.

61. Makarova,K.S., Haft,D.H., Barrangou,R., Brouns,S.J.J., Charpentier,E., Horvath,P., Moineau,S., Mojica,F.J.M., Wolf,Y.I., Yakunin,A.F., *et al.* (2011) Evolution and classification of the CRISPR-Cas systems. *Nat. Rev. Microbiol.*, **9**, 467–477.

62. Makarova,K.S., Wolf,Y.I., Alkhnbashi,O.S., Costa,F., Shah,S.A., Saunders,S.J., Barrangou,R., Brouns,S.J.J., Charpentier,E., Haft,D.H., *et al.* (2015) An updated evolutionary classification of CRISPR-Cas systems. *Nat. Rev. Microbiol.*, **13**, 722–736.

63. Peters,J.E., Makarova,K.S., Shmakov,S. and Koonin,E. V. (2017) Recruitment of CRISPR-Cas systems by Tn7-like transposons. *Proc. Natl. Acad. Sci. U. S. A.*, **114**, E7358–E7366.

64. Gleditzsch,D., Müller-Esparza,H., Pausch,P., Sharma,K., Dwarakanath,S., Urlaub,H., Bange,G. and Randau,L. (2016) Modulating the cascade architecture of a minimal type I-F CRISPR-Cas system. *Nucleic Acids Res.*, **44**, 5872–5882.

65. McDonald,N.D., Regmi,A., Morreale,D.P., Borowski,J.D. and Fidelma Boyd,E. (2019) CRISPR-Cas systems are present predominantly on mobile genetic elements in Vibrio species. *BMC Genomics*, **20**.

66. Gao,L., Altae-Tran,H., Böhning,F., Makarova,K.S., Segel,M., Schmid-Burgk,J.L., Koob,J., Wolf,Y.I., Koonin,E. V. and Zhang,F. (2020) Diverse enzymatic activities mediate antiviral immunity in prokaryotes. *Science*, **369**, 1077–1084.

67. Flemming,H.C. and Wuertz,S. (2019) Bacteria and archaea on Earth and their abundance in biofilms. *Nat. Rev. Microbiol.*, **17**, 247–260.

68. Granato,E.T., Meiller-Legrand,T.A. and Foster,K.R. (2019) The evolution and ecology of bacterial warfare. *Curr. Biol.*, **29**, R521–R537.

69. Peterson,S.B., Bertolli,S.K. and Mougous,J.D. (2020) The central role of interbacterial antagonism in bacterial fife. *Curr. Biol.*, **30**, R1203–R1214.

70. Cianfanelli,F.R., Monlezun,L. and Coulthurst,S.J. (2016) Aim, load, fire: the type VI secretion system, a bacterial nanoweapon. *Trends Microbiol.*, **24**, 51–62.

71. Taylor,N.M.I., van Raaij,M.J. and Leiman,P.G. (2018) Contractile injection systems of bacteriophages and related systems. *Mol. Microbiol.*, **108**, 6–15.

72. Galán,J.E. and Waksman,G. (2018) Protein-injection machines in bacteria. *Cell*, **172**, 1306–1318.

73. Boyer,F., Fichant,G., Berthod,J., Vandenbrouck,Y. and 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**.

74. Borgeaud,S., Metzger,L.C., Scrignari,T. and Blokesch,M. (2015) The type VI secretion system of *Vibrio cholerae* fosters horizontal gene transfer. *Science*, **347**, 63–67.

75. Metzger,L.C., Stutzmann,S., Scrignari,T., Van der Henst,C., Matthey,N. and Blokesch,M. (2016) Independent regulation of type VI secretion in *Vibrio cholerae* by TfoX and TfoY. *Cell Rep.*, **15**, 951–958.

76. Unterweger,D., Kitaoka,M., Miyata,S.T., Bachmann,V., Brooks,T.M., Moloney,J., Sosa,O., Silva,D., Duran-Gonzalez,J., Provenzano,D., *et al.* (2012) Constitutive type VI secretion system expression gives *Vibrio cholerae* intra- and interspecific competitive advantages. *PLoS One*, **7**, e48320.

77. Pukatzki,S., Ma,A.T., Sturtevant,D., Krastins,B., Sarracino,D., Nelson,W.C., Heidelberg,J.F. and Mekalanos,J.J. (2006) Identification of a conserved bacterial protein secretion system in *Vibrio cholerae* using the Dictyostelium host model system. *Proc. Natl. Acad. Sci. U. S. A.*, **103**, 1528–33.

78. Danka,E.S., Garcia,E.C. and Cotter,P.A. (2017) Are CDI systems multicolored, facultative, helping greenbeards? *Trends Microbiol.*, **25**, 391–401.

79. Ruhe,Z.C., Subramanian,P., Song,K., Nguyen,J.Y., Stevens,T.A., Low,D.A., Jensen,G.J. and Hayes,C.S. (2018) Programmed secretion arrest and receptor-triggered toxin export during antibacterial contact-dependent growth inhibition. *Cell*, **175**, 921–933.

80. Ruhe,Z.C., Low,D.A. and Hayes,C.S. (2013) Bacterial contact-dependent growth inhibition. *Trends Microbiol.*, **21**, 230–237.

81. Aoki,S.K., Diner,E.J., De Roodenbeke,C.T.K., Burgess,B.R., Poole,S.J., Braaten,B.A., Jones,A.M., Webb,J.S., Hayes,C.S., Cotter,P.A., *et al.* (2010) A widespread family of polymorphic contact-dependent toxin delivery systems in bacteria. *Nature*, **468**, 439–442.

82. Anderson,M.S., Garcia,E.C. and Cotter,P.A. (2012) The Burkholderia *bcpAIOB* genes define unique classes of two-partner secretion and contact dependent growth inhibition systems. *PLoS Genet.*, **8**, e1002877.

83. Garcia,E.C., Perault,A.I., Marlatt,S.A. and Cotter,P.A. (2016) Interbacterial signaling via *Burkholderia* contact dependent growth inhibition system proteins. *Proc. Natl. Acad. Sci. U. S. A.*, **113**, 8296–8301.

84. Aoki,S.K., Pamma,R., Hernday,A.D., Bickham,J.E., Braaten,B.A. and Low,D.A. (2005) Contact-dependent inhibition of growth in *Escherichia coli*. *Science*, **309**, 1245–1248.

85. Gonzalez,D. and Mavridou,D.A.I. (2019) Making the best of aggression: the many dimensions of bacterial toxin regulation. *Trends Microbiol.*, **27**, 897–905.

86. Willett,J.L.E., Gucinski,G.C., Fatherree,J.P., Low,D.A. and Hayes,C.S. (2015) Contact-dependent growth inhibition toxins exploit multiple independent cell-entry pathways. *Proc. Natl. Acad. Sci. U. S. A.*, **112**, 11341–11346.

87. Cascales,E., Buchanan,S.K., Duché,D., Kleanthous,C., Lloubès,R., Postle,K., Riley,M., Slatin,S. and Cavard,D. (2007) Colicin biology. *Microbiol. Mol. Biol. Rev.*, **71**, 158–229.

88. Ghequire,M.G.K. and De Mot,R. (2014) Ribosomally encoded antibacterial proteins and peptides from Pseudomonas. *FEMS Microbiol. Rev.*, **38**, 523–568.

89. Ghequire,M.G.K. and De Mot,R. (2015) The tailocin tale: peeling off phage tails. *Trends Microbiol.*, **23**, 587–590.

90. Michel-Briand,Y. and Baysse,C. (2002) The pyocins of *Pseudomonas aeruginosa*. *Biochimie*, **84**, 499–510.

91. Parret,A.H.A. and De Mot,R. (2002) Bacteria killing their own kind: novel bacteriocins of *Pseudomonas* and other γ-proteobacteria. *Trends Microbiol.*, **10**, 107–112.

92. Buist,G., Steen,A., Kok,J. and Kuipers,O.P. (2008) LysM, a widely distributed protein motif for binding to (peptido)glycans. *Mol. Microbiol.*, **68**, 838–847.

93. Salomon,D., Kinch,L.N., Trudgian,D.C., Guo,X., Klimko,J.A., Grishin,N. V., Mirzaei,H. and Orth,K. (2014) Marker for type VI secretion system effectors. *Proc. Natl. Acad. Sci.*, **111**, 9271–9276.

94. Maier,B. and Wong,G.C.L. (2015) How bacteria use type IV pili machinery on surfaces. *Trends Microbiol.*, **23**, 775–788.

95. Berry,J.L. and Pelicic,V. (2015) Exceptionally widespread nanomachines composed of type IV pilins: The prokaryotic Swiss Army knives. *FEMS Microbiol. Rev.*, **39**, 134–154.

96. Craig,L., Pique,M.E. and Tainer,J.A. (2004) Type IV pilus structure and bacterial pathogenicity. *Nat. Rev. Microbiol.*, **2**, 363–378.

97. Blokesch,M. (2014) The lifestyle of *Vibrio cholerae* fosters gene transfers. *Microbe*, **9**, 64–70.

98. Tamplin,M.L., Gauzens,A.L., Huq,A., Sack,D.A. and Colwell,R.R. (1990) Attachment of *Vibrio cholerae* serogroup O1 to zooplankton and phytoplankton of Bangladesh waters. *Appl. Environ. Microbiol.*, **56**, 1977–1980.

99. Meibom,K.L., Li,X.B., Nielsen,A.T., Wu,C.-Y., Roseman,S. and Schoolnik,G.K. (2004) The *Vibrio cholerae* chitin utilization program. *Proc. Natl. Acad. Sci.*, **101**, 2524–2529.

100. Meibom,K.L., Blokesch,M., Dolganov,N.A., Wu,C.Y. and Schoolnik,G.K. (2005) Chitin induces natural competence in *Vibrio cholerae*. *Science*, **310**, 1824–1827.

101. Seitz,P. and Blokesch,M. (2013) DNA-uptake machinery of naturally competent *Vibrio cholerae*. *Proc. Natl. Acad. Sci.*, **110**, 17987–17992.

102. Adams,D.W., Stutzmann,S., Stoudmann,C. and Blokesch,M. (2019) DNA-uptake pili of *Vibrio cholerae* are required for chitin colonization and capable of kin recognition via sequence-specific self-interaction. *Nat. Microbiol.*, **4**, 1545–1557.

103. Watnick,P.I. and Kolter,R. (1999) Steps in the development of a *Vibrio cholerae* El Tor biofilm. *Mol. Microbiol.*, **34**, 586–595.

104. Watnick,P.I., Fullner,K.J. and Kolter,R. (1999) A role for the mannose-sensitive hemagglutinin in biofilm formation by *Vibrio cholerae* El Tor. *J. Bacteriol.*, **181**, 3606–3609.

105. Chiavelli,D.A., Marsh,J.W. and Taylor,R.K. (2001) The mannose-sensitive hemagglutinin of *Vibrio cholerae* promotes adherence to zooplankton. *Appl. Environ. Microbiol.*, **67**, 3220–3225.

106. Utada,A.S., Bennett,R.R., Fong,J.C.N., Gibiansky,M.L., Yildiz,F.H., Golestanian,R. and Wong,G.C.L. (2014) *Vibrio cholerae* use pili and flagella synergistically to effect motility switching and conditional surface attachment. *Nat. Commun.*, **5**, 4913.

107. Hu,B., Khara,P. and Christie,P.J. (2019) Structural bases for F plasmid conjugation and F pilus biogenesis in *Escherichia coli*. *Proc. Natl. Acad. Sci. U. S. A.*, **116**, 14222–14227.

108. Lawley,T.D., Klimke,W.A., Gubbins,M.J. and Frost,L.S. (2003) F factor conjugation is a true type IV secretion system. *FEMS Microbiol. Lett.*, **224**, 1–15.

109. Chandran Darbari,V. and Waksman,G. (2015) Structural biology of bacterial type IV secretion systems. *Annu. Rev. Biochem.*, **84**, 603–629.

110. Matson,J.S., Withey,J.H. and DiRita,V.J. (2007) Regulatory networks controlling *Vibrio cholerae* virulence gene expression. *Infect. Immun.*, **75**, 5542–5549.

111. Schuhmacher,D.A. and Klose,K.E. (1999) Environmental signals modulate ToxT-dependent virulence factor expression in *Vibrio cholerae*. *J. Bacteriol.*, **181**, 1508–1514.

112. Dorman,M.J. and Dorman,C.J. (2018) Regulatory hierarchies controlling virulence gene expression in *Shigella flexneri* and *Vibrio cholerae*. *Front. Microbiol.*, **9**, 2686.

113. Midgett,C.R., Almagro-Moreno,S., Pellegrini,M., Taylor,R.K., Skorupski,K. and Kull,F.J. (2017) Bile salts and alkaline pH reciprocally modulate the interaction between the periplasmic domains of *Vibrio cholerae* ToxR and ToxS. *Mol. Microbiol.*, **105**, 258–272.

114. Peterson,K.M. and Gellings,P.S. (2018) Multiple intraintestinal signals coordinate the regulation of *Vibrio cholerae* virulence determinants. *Pathog. Dis.*, **76**, ftx126.

115. Ellison,C.K., Kan,J., Dillard,R.S., Kysela,D.T., Hampton,C.M., Ke,Z., Wright,E.R., Biais,N., Dalia,A.B. and Brun,Y. V. (2017) Obstruction of pilus retraction stimulates bacterial surface sensing. *Science*, **358**, 535–538.

116. Vesel,N. and Blokesch,M. (2021) Pilus production in *Acinetobacter baumannii* is growth phase dependent and essential for natural transformation. *J. Bacteriol.*, **203**, e00034-21.

117. Adams,D.W., Pereira,J.M., Stoudmann,C., Stutzmann,S. and Blokesch,M. (2019) The type IV pilus protein PilU functions as a PilT-dependent retraction ATPase. *PLoS Genet.*, **15**, e1008393.

118. Denise,R., Abby,S.S. and Rocha,E.P.C. (2019) Diversification of the type IV filament superfamily into machines for adhesion, protein secretion, DNA uptake, and motility. *PLoS Biol.,* **17**, e3000390.

119. Tomich,M., Planet,P.J. and Figurski,D.H. (2007) The *tad* locus: Postcards from the widespread colonization island. *Nat. Rev. Microbiol.*, **5**, 363–375.

120. Giltner,C.L., Nguyen,Y. and Burrows,L.L. (2012) Type IV Pilin Proteins: Versatile Molecular Modules. *Microbiol. Mol. Biol. Rev.*, **76**, 740–772.

121. Planet,P.J., Kachlany,S.C., Fine,D.H., DeSalle,R. and Figurski,D.H. (2003) The widespread colonization island of *Actinobacillus actinomycetemcomitans*. *Nat. Genet.*, **34**, 193–198.

122. Pu,M., Duriez,P., Arazi,M. and Rowe-Magnus,D.A. (2018) A conserved tad pilus promotes *Vibrio vulnificus* oyster colonization. *Environ. Microbiol.*, **20**, 828–841.

123. Pu,M. and Rowe-Magnus,D.A. (2018) A Tad pilus promotes the establishment and resistance of *Vibrio vulnificus* biofilms to mechanical clearance. *npj Biofilms Microbiomes*, **4**.

124. Mukherjee,A., Mammel,M.K., LeClerc,J.E. and Cebula,T.A. (2008) Altered utilization of N-acetyl-D-galactosamine by *Escherichia coli* O157:H7 from the 2006 spinach outbreak. *J. Bacteriol.*, **190**, 1710–1717.

125. Brinkkötter,A., Klöß,H., Alpert,C.A. and Lengeler,J.W. (2000) Pathways for the utilization of N-acetyl-galactosamine and galactosamine in *Escherichia coli*. *Mol. Microbiol.*, **37**, 125–135.

126. Leyn,S.A., Gao,F., Yang,C. and Rodionov,D.A. (2012) N-acetylgalactosamine utilization pathway and regulon in proteobacteria: genomic reconstruction and experimental characterization in *Shewanella*. *J. Biol. Chem.*, **287**, 28047–28056.

127. Payne,S.M., Mey,A.R. and Wyckoff,E.E. (2016) *Vibrio* iron transport: evolutionary adaptation to life in multiple environments. *Microbiol. Mol. Biol. Rev.*, **80**, 69–90.

128. Thode,S.K., Rojek,E., Kozlowski,M., Ahmad,R. and Haugen,P. (2018) Distribution of siderophore gene systems on a Vibrionaceae phylogeny: database searches, phylogenetic analyses and evolutionary perspectives. *PLoS One*, **13**, e0191860.

129. Neilands,J.B. (1995) Siderophores: structure and function of microbial iron transport compounds. *J. Biol. Chem.*, **270**, 26723–26726.

130. Osorio,C.R., Juiz-Río,S. and Lemos,M.L. (2006) A siderophore biosynthesis gene cluster from the fish pathogen *Photobacterium damselae* subsp. *piscicida* is structurally and functionally related to the Yersinia high-pathogenicity island. *Microbiology*, **152**, 3327–3341.

131. Osorio,C.R., Rivas,A.J., Balado,M., Fuentes-Monteverde,J.C., Rodríguez,J., Jiménez,C., Lemos,M.L. and Waldor,M.K. (2015) A transmissible plasmid-borne pathogenicity island confers piscibactin biosynthesis in the fish pathogen *Photobacterium damselae* subsp. *piscicida*. *Appl. Environ. Microbiol.*, **81**, 5867–5879.

132. Sleator,R.D. and Hill,C. (2002) Bacterial osmoadaptation: the role of osmolytes in bacterial stress and virulence. *FEMS Microbiol. Rev.*, **26**, 49–71.

133. Kempf,B. and Bremer,E. (1998) Uptake and synthesis of compatible solutes as microbial stress responses to high-osmolality environments. *Arch. Microbiol.*, **170**, 319–330.

134. Kapfhammer,D., Karatan,E., Pflughoeft,K.J. and Watnick,P.I. (2005) Role for glycine betaine transport in *Vibrio cholerae* osmoadaptation and biofilm formation within microbial communities. *Appl. Environ. Microbiol.*, **71**, 3840–3847.

135. Naughton,L.M., Blumerman,S.L., Carlberg,M. and Boyd,E.F. (2009) Osmoadaptation among *Vibrio* species and unique genomic features and physiological responses of *Vibrio parahaemolyticus*. *Appl. Environ. Microbiol.*, **75**, 2802–2810.

136. Wen,Y., Kim,I.H., Son,J.S., Lee,B.H. and Kim,K.S. (2012) Iron and quorum sensing coordinately regulate the expression of vulnibactin biosynthesis in *Vibrio vulnificus*. *J. Biol. Chem.*, **287**, 26727–26739.

137. McRose,D.L., Baars,O., Seyedsayamdost,M.R. and Morel,F.M.M. (2018) Quorum sensing and iron regulate a two-for-one siderophore gene cluster in *Vibrio harveyi*. *Proc. Natl. Acad. Sci. U. S. A.*, **115**, 7581–7586.

138. van Kessel,J.C., Rutherford,S.T., Cong,J.P., Quinodoz,S., Healy,J. and Basslera,B.L. (2015) Quorum sensing regulates the osmotic stress response in *Vibrio harveyi*. *J. Bacteriol.*, **197**, 73–80.

139. Widder,E.A. (2010) Bioluminescence in the ocean: origins of biological, chemical and ecological diversity. *Science*, **328**, 704–708.

140. Waidmann,M.S., Bleichrodt,F.S., Laslo,T. and Riedel,C.U. (2011) Bacterial luciferase reporters the swiss army knife of molecular biology. *Bioeng. Bugs*, **2**, 8–16.

141. Urbanczyk,H., Ast,J.C., Kaeding,A.J., Oliver,J.D. and Dunlap,P. V. (2008) Phylogenetic analysis of the incidence of lux gene horizontal transfer in Vibrionaceae. *J. Bacteriol.*, **190**, 3494–3504.

142. Brodl,E., Winkler,A. and Macheroux,P. (2018) Molecular mechanisms of bacterial bioluminescence. *Comput. Struct. Biotechnol. J.*, **16**, 551–564.

143. Grim,C.J., Taviani,E., Alam,M., Huq,A., Sack,R.B. and Colwell,R.R. (2008) Occurrence and expression of luminescence in *Vibrio cholerae*. *Appl. Environ. Microbiol.*, **74**, 708–715.

144. Naim,R., Yanagihara,I., Iida,T. and Honda,T. (2001) Vibrio parahaemolyticus thermostable direct hemolysin can induce an apoptotic cell death in Rat-1 cells from inside and ouside of the cells. *FEMS Microbiol. Lett.*, **195**, 237–244.

145. Raghunath,P. (2014) Roles of thermostable direct hemolysin (TDH) and TDH-related hemolysin (TRH) in Vibrio parahaemolyticus. *Front. Microbiol.*, **5**, 2010–2013.

146. Miller,K.A., Tomberlin,K.F. and Dziejman,M. (2019) Vibrio variations on a type three theme. *Curr. Opin. Microbiol.*, **47**, 66–73.

147. Abby,S.S. and Rocha,E.P.C. (2012) The non-flagellar type III secretion system evolved from the bacterial flagellum and diversified into host-cell adapted systems. *PLoS Genet.*, **8**, e1002983.

148. Galán,J.E. (2009) Common themes in the design and function of bacterial effectors. *Cell Host Microbe*, **5**, 571–579.

149. Buttner,D. (2012) Protein export according to schedule: architecture, assembly, and regulation of type III secretion systems from plant- and animal-pathogenic bacteria. *Microbiol. Mol. Biol. Rev.*, **76**, 262–310.

150. Deng,W., Marshall,N.C., Rowland,J.L., McCoy,J.M., Worrall,L.J., Santos,A.S., Strynadka,N.C.J. and Finlay,B.B. (2017) Assembly, structure, function and regulation of type III secretion systems. *Nat. Rev. Microbiol.*, **15**, 323–337.

151. Wagner,S., Grin,I., Malmsheimer,S., Singh,N., Torres-Vargas,C.E. and Westerhausen,S. (2018) Bacterial type III secretion systems: a complex device for the delivery of bacterial effector proteins into

eukaryotic host cells. *FEMS Microbiol. Lett.*, **365**, fny201.

152. Lara-Tejero,M. and Galán,J.E. (2019) The injectisome, a complex nanomachine for protein injection into mammalian cells. *EcoSal Plus*, **8**, 10.1128.

153. Dziejman,M., Serruto,D., Tam,V.C., Sturtevant,D., Diraphat,P., Faruque,S.M., Rahman,M.H., Heidelberg,J.F., Decker,J., Li,L., *et al.* (2005) Genomic characterization of non-O1, non-O139 *Vibrio cholerae* reveals genes for a type III secretion system. *Proc. Natl. Acad. Sci.*, **102**, 3465–3470.

154. Makino,K., Oshima,K., Kurokawa,K., Yokoyama,K., Uda,T., Tagomori,K., Iijima,Y., Najima,M., Nakano,M., Yamashita,A., *et al.* (2003) Genome sequence of *Vibrio parahaemolyticus*: A pathogenic mechanism distinct from that of *V. cholerae. Lancet*, **361**, 743–749.

155. Tam,V.C., Serruto,D., Dziejman,M., Brieher,W. and Mekalanos,J.J. (2007) A Type III secretion system in *Vibrio cholerae* translocates a formin/spire hybrid-like actin nucleator to promote intestinal colonization. *Cell Host Microbe*, **1**, 95–107.

156. Shin,O.S., Tam,V.C., Suzuki,M., Ritchie,J.M., Bronson,R.T., Waldor,M.K. and Mekalanos,J.J. (2011) Type III secretion is essential for the rapidly fatal diarrheal disease caused by non-O1, non-O139 *Vibrio cholerae*. *MBio*, **2**, e00106-11.

157. Darling,A.E., Mau,B. and Perna,N.T. (2010) Progressivemauve: multiple genome alignment with gene gain, loss and rearrangement. *PLoS One*, **5**, e11147.

158. Hueck,C.J. (1998) Type III protein secretion systems in bacterial pathogens of animals and plants. *Microbiol. Mol. Biol. Rev.*, **62**, 379–433.

159. Alam,A., Miller,K.A., Chaand,M., Butler,J.S. and Dziejman,M. (2011) Identification of *Vibrio cholerae* type III secretion system effector proteins. *Infect. Immun.*, **79**, 1728–1740.

160. Rosqvist,R., Forsberg,A., Rimpilainen,M., Bergman,T. and Wolf-Watz,H. (1990) The cytotoxic protein YopE of *Yersinia* obstructs the primary host defence. *Mol. Microbiol.*, **4**, 657–667.

161. Tam,V.C., Suzuki,M., Coughlin,M., Saslowsky,D., Biswas,K., Lencer,W.I., Faruque,S.M. and Mekalanos,J.J. (2010) Functional analysis of VopF activity required for colonization in *Vibrio cholerae*. *MBio*, **1**, e00289-10.

162. Suzuki,M., Danilchanka,O. and Mekalanos,J.J. (2014) *Vibrio cholerae* T3SS effector VopE modulates mitochondrial dynamics and innate immune signaling by targeting Miro GTPases. *Cell Host Microbe*, **16**, 581–591.

163. Alam,A., Tam,V., Hamilton,E. and Dziejman,M. (2010) vttRa and vttRB encode ToxR family proteins that mediate bile-induced expression of type three secretion system genes in a non-O1/non-O139 *Vibrio cholerae* strain. *Infect. Immun.*, **78**, 2554–2570.

164. Ellermeier,J.R. and Slauch,J.M. (2007) Adaptation to the host environment: regulation of the SPI1 type III secretion system in *Salmonella enterica* serovar Typhimurium. *Curr. Opin. Microbiol.*, **10**, 24–29.

165. Salomon,D., Guo,Y., Kinch,L.N., Grishin,N. V., Gardner,K.H. and Orth,K. (2013) Effectors of animal and plant pathogens use a common domain to bind host phosphoinositides. *Nat. Commun.*, **4**, 2973.

166. Kaniga,K., Trollinger,D. and Galan,J.E. (1995) Identification of two targets of the type III protein secretion system encoded by the inv and spa loci of *Salmonella typhimurium* that have homology to the *Shigella* IpaD and IpaA proteins. *J. Bacteriol.*, **177**, 7078–7085.

167. Miller,K.A., Chaand,M., Gregoire,S., Yoshida,T., Beck,L.A., Ivanov,A.I. and Dziejman,M. (2016) Characterization of *V. cholerae* T3SS-dependent cytotoxicity in cultured intestinal epithelial cells. **18**, 1857–1870.

168. Calder,T., de Souza Santos,M., Attah,V., Klimko,J., Fernandez,J., Salomon,D., Krachler,A.M. and Orth,K. (2014) Structural and regulatory mutations in *Vibrio parahaemolyticus* type III secretion systems display variable effects on virulence. *FEMS Microbiol. Lett.*, **361**, 107–114.

169. Lo Scrudato,M. and Blokesch,M. (2012) The regulatory network of natural competence and transformation of *Vibrio cholerae*. *PLoS Genet.*, **8**, e1002778.

170. Langmead,B. and Salzberg,S.L. (2012) Fast gapped-read alignment with Bowtie 2. *Nat. Methods*, **9**, 357–359.

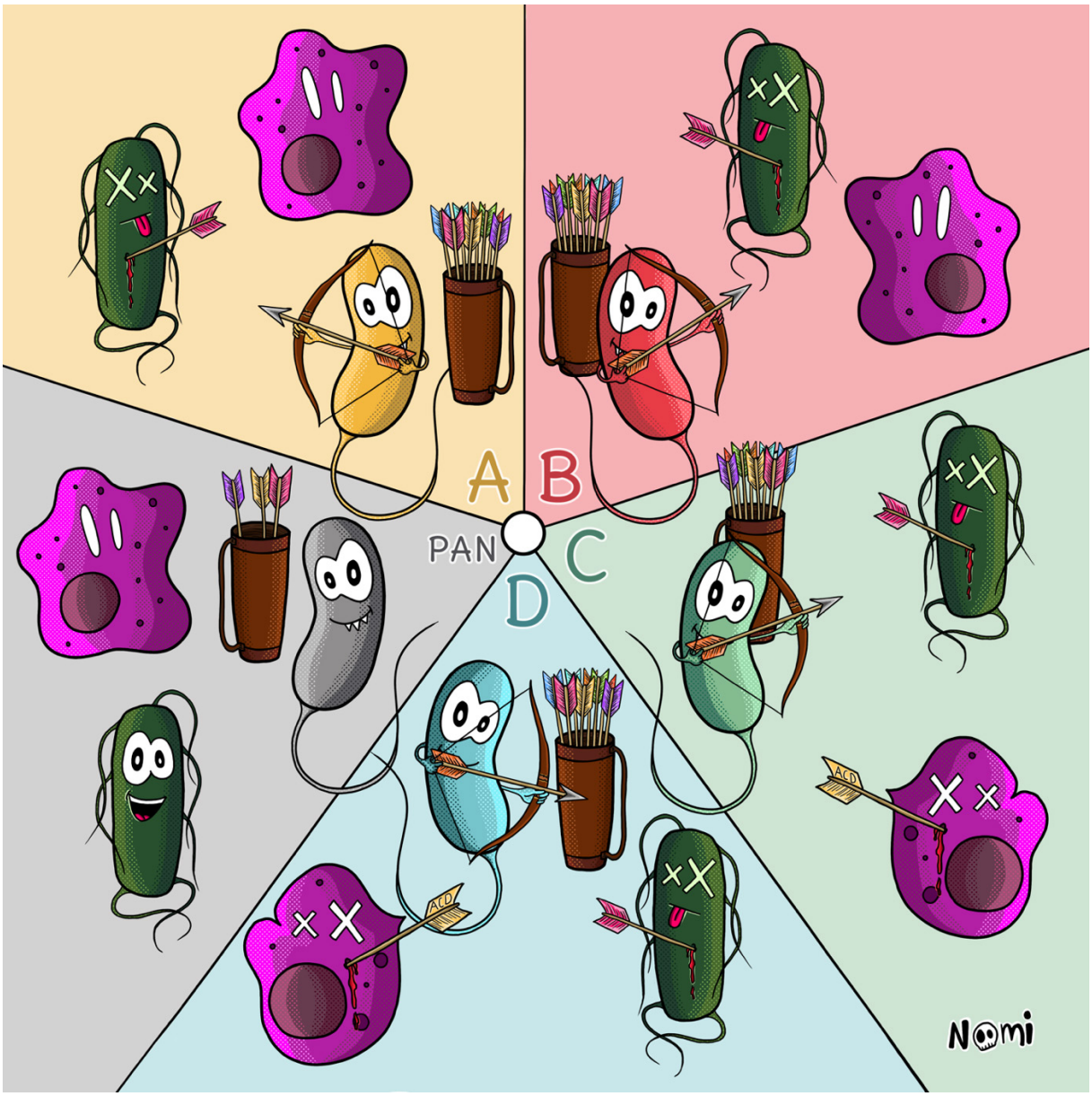
171. Anders,S., Pyl,P.T. and Huber,W. (2015) HTSeq-A Python framework to work with high-throughput sequencing data. *Bioinformatics*, **31**, 166–169.

172. Love,M.I., Huber,W. and Anders,S. (2014) Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol.*, **15**.

173. Maaten Van Der,L. (2014) Accelerating t-SNE using tree-based algorithms. *J. Mach. Learn. Res.*, **15**, 3221–3245.

174. Yildiz,F.H. and Schoolnik,G.K. (1998) Role of *rpoS* in stress survival and virulence of *Vibrio cholerae*. *J. Bacteriol.*, **180**, 773–784.

4 INTERBACTERIAL COMPETITION AND ANTI-PREDATORY BEHAVIOR OF ENVIRONMENTAL *VIBRIO CHOLERAE* STRAINS



Drawing by Noémie Matthey, PhD

This chapter is a post-print version of an article published as:

N.C. Drebes Dörr and M. Blokesch*. Interbacterial competition and anti-predatory behavior of environmental *Vibrio cholerae* strains. Environ Microbiology (2020) 22(10), 4485-4504. doi:10.1111/1462-2920.15224

Laboratory of Molecular Microbiology, Global Health Institute, School of Life Sciences, École Polytechnique Fédérale de Lausanne (EPFL), CH-1015 Lausanne, Switzerland.

*Correspondence: Melanie Blokesch; EPFL-SV-UPBLO, Station 19, École Polytechnique Fédérale de Lausanne (EPFL), CH-1015 Lausanne, Switzerland; Phone: +41 21 693 0653; Email: melanie.blokesch@epfl.ch.

Running title: Toxicity of environmental *Vibrio cholerae* strains

Keywords: type VI secretion, hemolysin, *Vibrio cholerae,* amoebae

4.1 ORIGINALITY-SIGNIFICANCE STATEMENT

This work contributes to the understanding of phenotypic consequences that differentiate diverse *Vibrio cholerae* strains. We focused on the type VI secretion system (T6SS) and the pore forming toxin hemolysin, which are tightly regulated in pandemic strains but remain constitutively active in non-pandemic isolates. We unveiled diverse arrays of T6SS effector/immunity modules in a set of environmental strains by long-read whole genome sequencing and *de novo* assembly. These modules determine whether the strains are able to evade amoebal predation and dictate their level of compatibility or competitiveness with one another.

4.2 SUMMARY

Vibrio cholerae isolates responsible for cholera pandemics represent only a small portion of the diverse strains belonging to this species. Indeed, most *V. cholerae* are encountered in aquatic environments. To better understand the emergence of pandemic lineages, it is crucial to discern what differentiates pandemic strains from their environmental relatives. Here, we studied the interaction of environmental *V. cholerae* with eukaryotic predators or competing bacteria and tested the contributions of the hemolysin and the type VI secretion system (T6SS) to those interactions. Both of these molecular weapons are constitutively active in environmental isolates but subject to tight regulation in the pandemic clade. We showed that several environmental isolates resist amoebal grazing and that this anti-grazing defense relies on the strains' T6SS and its actin-cross-linking domain (ACD)-containing tip protein. Strains lacking the ACD were unable to defend themselves against grazing amoebae but maintained high levels of T6SS-dependent interbacterial killing. We explored the latter phenotype through whole-genome sequencing of fourteen isolates, which unveiled a wide array of novel T6SS effector and (orphan) immunity proteins. By combining these *in*

silico predictions with experimental validations, we showed that highly similar but nonidentical immunity proteins were insufficient to provide cross-immunity among those wild strains.

4.3 INTRODUCTION

Diarrheal diseases can be caused by a variety of microorganisms, including the causative agent of cholera, *Vibrio cholerae*, which infects up to 4 million people every year (1). Cholera often spreads from its endemic area around the Ganges delta and has reached almost the entire world in the reported seven pandemics that have been witnessed since 1817. In addition to these pandemics, important localized outbreaks have occurred over the years, especially following natural disasters (2–4).

V. cholerae strains can be classified into serogroups based on more than 200 different O-antigens. The O1 serogroup is considered the primary cause of previous (e.g., classical serotype) and ongoing (e.g., El Tor serotype) cholera pandemics (2, 5). Isolates belonging to the O139 serogroup are genetically related to the 7th pandemic O1 El Tor strains (6) but are rarely associated with disease outbreaks currently (2, 4). Even though this species is best known due to the life-threatening disease it causes, the vast majority of *V. cholerae* are common members of aquatic habitats. These mostly non-O1/non-O139 serogroup strains are thought to frequently associate with zooplankton and shellfish, and *V. cholerae* uses their chitinous exoskeletons as a source of carbon and nitrogen (2, 5, 7). Environmental isolates are typically considered harmless to humans despite reported associations with mild to severe forms of diarrhea (8–14).

A major question, not only in the cholera field but also for infectious diseases in general, is how pathogenic isolates evolve from their non-pathogenic environmental

progenitors. In the case of *V. cholerae*, pandemic patient isolates are remarkably clonal, in sharp contrast to the high genomic variability encountered in environmental isolates (2, 15–20). Two major genetic features common to all pandemic strains are the CTXΦ prophage and the *Vibrio* pathogenicity island (VPI-1 or TCP island). These genetic elements harbor genes encoding the main virulence factors, e.g., cholera toxin (CTX) and toxin-coregulated pilus (TCP) (21, 22). The possession of these genomic regions, however, is not unique nor a 'deterministic factor' of pandemic strains. In fact, *V. cholerae* samples collected in several regions of the world have exposed environmental CTX- and/or TCP-positive strains (23–28). Additionally, *V. cholerae* strains that have caused localized cholera outbreaks without reaching pandemic levels have been reported, such as the O37 serogroup strains V52 and ATCC25872 (5, 16, 29, 30). Therefore, the terms "toxigenic" (ability to cause cholera) and "pandemic" (here, current 7th pandemic-causing O1 El Tor strains) are not synonymous. What exactly determines whether strains become pandemic or not is still not fully understood and probably involves diverse aspects ranging from genetic content and strain-specific phenotypes, along with the location of the initial outbreaks and ultimately towards social and sanitary factors, to name a few. Pertinent to the bacteria-related aspects, previous studies hypothesized the existence of virulence adaptive polymorphisms (VAP) circulating in environmental strains. These VAPs were suggested to be a prerequisite for pandemic transition before the horizontal acquisition of CTXΦ and VPI-1 (24).

To better understand their emergence, the differences between pandemic strains and their environmental relatives need to be deciphered. Such differences most likely include phenotypic alterations that are not easily predictable by genomics. In the context of phenotypic variation, two minor virulence factors, namely, the type VI secretion system (T6SS) and the pore-forming toxin hemolysin, are of special interest as they are differentially produced in pandemic compared with non-pandemic toxigenic or environmental strains. The

T6SS is a molecular killing device that resembles an inverted contractile bacteriophage tail and it is present in approximately 25% of all Gram-negative bacteria (31–34). It is composed of a membrane-spanning portion, a tube structure made of stacks of Hcp hexamer rings, and a tube-surrounding sheath. The sheath is composed of the two proteins VipA and VipB which, upon contraction, propels the inner tube out of the cell together with its effector-decorated tip proteins (VgrG and PAAR) (35–37). Secreted Hcp therefore serves as an indicator of T6SS activity (38–41). The secreted effector proteins mostly target conserved cellular components, such as membranes, bacterial peptidoglycan, nucleic acids, or the eukaryotic cytoskeleton (42, 43). T6SS effectors can be either secreted as (i) cargos that interact with T6SS structural proteins, such as Hcp or the tip protein VgrG; or as (ii) C-terminal extensions of VgrG, Hcp, or PAAR (so-called “evolved” proteins; (38, 44, 45)). Notably, T6SS-producing bacteria protect themselves against their toxic effector repertoire by the production of effector-cognate immunity proteins. These effector/immunity (E/I) pairs are usually encoded adjacent to each other (46–49).

In current pandemic strains, the T6SS is encoded by four gene clusters: the large cluster and three auxiliary clusters. The large cluster primarily codes for structural proteins (recently reviewed by (50)), including the evolved tip protein VgrG3, which possesses a C-terminal lysozyme-like domain for peptidoglycan degradation (48, 51). Auxiliary clusters 1 and 2 are both composed of genes encoding Hcp and VgrG (evolved in cluster 1 and structural in cluster 2), an adaptor protein (Tap1 or VasW; (52, 53)) and an E/I module. The auxiliary cluster 1 effector TseL is a bifunctional lipase with anti-bacterial and anti-eukaryotic activity (47, 51, 54). The auxiliary cluster 2 effector VasX acts as a pore-forming toxin due to its bacterial/eukaryotic colicin-like membrane-disrupting activity (43, 55, 56). Finally, the T6SS auxiliary cluster 3 is composed of genes encoding a second copy of PAAR (a tip-sharpening protein that extends from VgrG; the first gene copy is at the start of the large

cluster; (57)) and a single E/I pair. When discovered, the aux 3 effector TseH was predicted to contain a hydrolase domain (58), while the recently reported crystal structure supported its role as a papain-like NlpC/P60 peptidase (59) with structural similarity to the T6SS effector Tse1 of *Pseudomonas aeruginosa* that contains bacteriolytic peptidoglycan amidase activity (60).

Another toxic protein that shows differential activity in pandemic compared with non-pandemic *V. cholerae* strains is the hemolysin protein (HlyA), which is likewise widespread among *Vibrio* species (61). HlyA is considered a minor virulence factor as it contributes to toxicity in the context of intestinal infections (62, 63). We previously showed that the proper timing of HlyA activity is a prerequisite for pandemic *V. cholerae* to establish a replication niche within the aquatic amoeba *Acanthamoeba castellanii*, while constitutive activity kills this host prematurely (64).

In this study, we deciphered phenotypic and genotypic differences between O1 El Tor pandemic strains and 15 environmental *V. cholerae* isolates. With respect to phenotypes, we focused primarily on T6SS- and hemolysin-specific outcomes exerted on competing bacteria or amoebal predators. We show that clade-specific anti-amoebal toxicity is dependent on the actin-cross-linking domain of a subset of T6SS effector proteins and that HlyA does not harm these predators under the tested conditions. Anti-bacterial activity, on the other hand, is widespread among the environmental isolates and is based on their constitutive T6SS production, which is a major phenotypic difference compared with the exquisitely regulated T6SS of pandemic *V. cholerae*. We also performed long-read PacBio whole-genome sequencing of the environmental isolates, which allowed us to perform basic comparative genomic analyses. Through this approach, we identified T6SS E/I modules as well as orphan immunity loci. Finally, we experimentally tested how the environmental

isolates compete with one another and how this interbacterial competition correlates with their T6SS E/I repertoire.

4.4 RESULTS AND DISCUSSION

4.4.1 Genome sequencing of environmental *V. cholerae* strains

Apart from the major virulence factors, previous observations suggested important phenotypic differences between pandemic and environmental *V. cholerae* strains. We therefore decided to study fifteen environmental isolates (Supplementary table 4.1) from diverse habitats along the central California coast, a region that is free of endemic or epidemic cholera. These strains were initially isolated in 2004 by Keymer and colleagues followed by basic characterization and comparative genome hybridization (CGH) analyses (65, 66). The latter approach was based on amplicon microarrays (representing 3,357 of 3,891 annotated open reading frames), which were designed using the first published *V. cholerae* genome sequence as a template, namely, pandemic O1 El Tor strain N16961 (67). Based on the presence or absence of the microarray-templated genes, the strains were classified into four clades (A-D) (65, 66), a classification we maintained throughout the current report. To better understand the accessory genome, including the T6SS E/I modules and to identify those genes that are novel when compared with strain N16961, we first whole-genome sequenced these strains using a long-read PacBio approach followed by the *de novo* assembly of their genomes. As a representative strain of the 7th cholera pandemic, we used strain O1 El Tor A1552 throughout this study (68). This strain is connected to a cholera outbreak in Peru in the 1990s. We recently reported its genome sequence, including more than 1,000 manual gene annotations, according to previous experimental validations (69).

Supplementary table 4.2 shows the sequencing details and the features of the closed genomes of the 15 environmental isolates. All genomes showed the dual chromosome architecture that is common for *Vibrio* species (70), a similar overall size of the two chromosomes, and average GC percentages within the same range as the one observed for the pandemic *V. cholerae* strains N16961 and A1552 (67, 69). The assembly pipeline also predicted megaplasmids of ~300kbp and 80kbp for four strains belonging to clades C and D, respectively. It should be noted that due to the size selection of the prepared sequencing libraries, putative smaller plasmids remained unidentified. Interestingly, we observed that the genomes of strains W6G and W7G were almost identical. This reflects the previous report by Keymer *et al*. In their original sampling study, the authors claimed that, based on CGH, 30 unique genotypes were identified within their collection of 41 environmental strains, while several genotypes were sampled multiple times from distinct sampling events. Indeed, upon direct comparison of the W6G and W7G genomes, we observed pairwise identities of 99.98%, 99.997% and 99.998% for chromosome 1, chromosome 2, and the megaplasmid, respectively. Both strains possessed the same genes in all three replicons, and the few observed differences were primarily single nucleotides indels (which might, in part, reflect sequencing artifacts). These data suggest almost clonality between these two environmental samples (W6G and W7G). In addition, our whole-genome sequencing data showed that strains E7G and SA7G of clade D also had high levels of identity (99.4%, 98.5%, and 99.9% pairwise identity of chromosomes 1, 2, and the megaplasmid, respectively), which again confirmed the previous CGH data (65). Besides SNPs and indels, these strains differ by the presence/absence of small genomic islands on both chromosomes and an additional, second megaplasmid in strain SA7G, which is missing in E7G. The majority of other strains differed more significantly and therefore allowed us to test links between specific phenotypes and the corresponding genotypes.

4.4.2 Only a subset of environmental isolates block amoebal predation

We initially wondered how these environmental isolates would behave when confronted with predatory grazers, such as bacterivorous amoebae. We therefore tested the representative 7th pandemic strain A1552 and the above-described collection of environmental strains for their ability to defend themselves against the grazing soil amoeba *Dictyostelium discoideum*. As shown in Figure 4.1a, a clade-specific behavior was observable in which strains from clade A and B appeared as non-toxic to *D. discoideum* and were efficiently grazed on while clade C and D strains completely resisted amoebal grazing. We concluded that anti-amoebal defenses significantly vary among the different environmental *V. cholerae* isolates.

4.4.3 Enhanced hemolysin production in environmental *V. cholerae* does not contribute to their anti-amoebal behavior

As we observed that a subset of the environmental strains had a strong anti-amoebal effect, we wondered whether this grazing inhibition was linked to toxic effectors of *V. cholerae*. Our group had previously demonstrated that the proper timing of the production or activity of the pore-forming toxin hemolysin HlyA was essential for pandemic *V. cholerae* to form a replication niche inside the aquatic amoebae *A. castellanii*. Indeed, while pandemic *V. cholerae* exerted a tight regulation over this toxin and thereby successfully infected the amoebae's contractile vacuole, constitutive hemolysin activity by the environmental isolates killed the host prematurely, preventing the formation of this replication niche (64). We speculated that hemolysin activity could also be involved in the resistance against *D. discoideum* grazing that we observed for the environmental strains from clades C and D. To follow up on this hypothesis, we first tested all environmental WT strains for hemolysis on blood agar plates. As shown in Figure 4.1b, all isolates had strong hemolytic activity, especially when compared with the pandemic strain A1552. To ensure that the hemolysis

was indeed caused by HlyA's activity, we interrupted the *hlyA* gene (loci comparable to locus tag VCA0219 in reference strain N16961) in a subset of the environmental strains through the integration of an antibiotic resistance cassette (the wild-type (WT) versions of the environmental strains as well as their respective mutants are listed in Supplementary table 4.1). The selection of this subset of environmental isolates was thereby based on two criteria: i) to represent each clade; and ii) to select those strains that showed efficient chitin-induced natural transformability, which allowed for genetic manipulation of those strains. For these representative strains, we confirmed that the observed blood cell lysis was indeed caused by the specific hemolysin HlyA (Figure 4.1c). After this confirmation, we tested the strains in the amoebal grazing assay. However, to our surprise, the hemolysin-deficient mutants behaved the same way as their parental WT strains, indicating that HlyA is not responsible for the strong anti-amoebal behavior that we observed for the clade C and D strains (Figure 4.1d). These results therefore suggest that, at least under the tested conditions, HlyA does not play an important role in anti-amoebal grazing defense.

4.4.4 Environmental isolates constitutively produce T6SS

As HlyA activity was ruled out as a defense mechanism, we moved on to check the potential involvement of other molecular weapons. We next considered the T6SS, due to its widespread occurrence in Gram-negative bacteria. Notably, *V. cholerae*'s T6SS was initially discovered due to its anti-eukaryotic activity that allowed the non-pandemic *V. cholerae* strain V52 to avoid predation by *D. discoideum* (38). Indeed, previous studies had indicated that, in contrast to the 7th pandemic strains, non-pandemic toxigenic *V. cholerae* (such as the two O37 serogroup strains V52 and ATCC25872; (38, 39, 64) as well as environmental isolates (28, 50, 71) maintain constitutive T6SS activity. This is in contrast to the silenced T6SS of pandemic strains, which are primarily induced by chitin or low c-di-GMP levels,

concomitantly with natural competence and the production of extracellular enzymes, respectively (72–76). To check whether the T6SS protects clade C and D strains, we first tested the general T6SS activity of the environmental isolates. As shown in Figure 4.2a, we observed that the environmental strains efficiently eradicated *Escherichia coli* prey bacteria. Only strain SA3G of clade B reproducibly killed prey with a reduced efficiency, even though residual T6SS activity was still observed when compared with the nonkilling pandemic strain A1552 (Figure 4.2a). However, these data alone do not unambiguously show whether the observed prey effacement was indeed T6SS-dependent or was instead the result of any other modes of interbacterial competition, such as contact-dependent inhibition, toxin secretion, bacteriocins, etc. (77–79). We therefore deleted the T6SS sheath protein-encoding gene *vipA* in each of the clade-representing strains and confirmed their lost T6SS activity by scoring for Hcp secretion. As illustrated in Figure 4.2b, all strains were able to produce Hcp protein but only the T6SS-active parental environmental strains were able to also secrete this protein into the supernatant. These data are therefore in agreement with the idea that the strains' T6SS is indeed constitutively active (at least under the tested conditions) and is the reason behind the observed interbacterial killing phenotype (Figure 4.2a). We confirmed the latter idea by comparing the WT and *vipA*-minus derivatives' killing ability in an interbacterial competition assay using *E. coli* as prey (Figure 4.2c).

4.4.5 Environmental strains use their VgrG-linked ACD of the T6SS to fight amoebae

Having recognized that all environmental strains constitutively produce their T6SS, we moved on to assess the involvement of this machinery in the anti-amoebal defense of clade C and D strains. Indeed, the observed inhibition of amoebal plaque formation by a subset of the environmental isolates was consistent with previous work by Unterweger and colleagues (71). These authors had studied four environmental isolates from estuaries of the Rio

Grande delta for anti-amoebal and anti-bacterial activity and observed that two of these isolates could not resist amoebal predation. These strains were, however, also unable to kill *E. coli* prey, and the reason for this interbacterial noncompetitiveness was a frameshift mutation in the intermediate T6SS regulatory protein-encoding gene *vasH* (38, 71). In contrast, apart from one exception (strain SA3G), all of the tested environmental isolates in our study efficiently eradicated *E. coli* prey (Figure 4.2a), indicating that the T6SS was, in general, functional and active. We therefore reassessed the amoebal plaque formation against the genetically modified T6SS mutant strains. As shown in Figure 4.2d, clade C and D isolates, whose parental WT strains completely blocked amoebal predation, became nontoxic when their T6SS was inactivated, indicating that their anti-amoebal defense was indeed linked to the T6SS and uniquely caused by the latter.

Considering that all of the environmental isolates have a constitutively active T6SS under the tested conditions and efficiently killed other bacteria (Figure 4.2a), we wondered why only the strains from clades C and D were able to use their T6SS as an anti-eukaryotic defense tool. To answer this question and to also characterize the full E/I modules of these strains, we inspected the T6SS clusters in the new genomic sequencing data and observed a clear clade specificity. Only those strains belonging to clades C and D encoded evolved VgrG1 proteins with a C-terminal actin cross-linking domain (ACD) (Figure 4.2 and Figure 4.3; Supplementary table 4.3). Strains from clades A and B, on the other hand, encoded only structurally relevant VgrG1 proteins without an evolved effector domain. Importantly, pandemic strains also encoded such an ACD as part of *vgrG1*, but, as noted above, these strains do not produce functional T6SSs without specific environmental cues and therefore show neither anti-bacterial (Figure 4.2a) nor anti-amoebal behavior (Figure 4.2d) under the tested conditions.

The ACD of VgrG1 in non-pandemic but toxigenic O37 serogroup strain V52 (16), which produces its T6SS constitutively, was previously shown to be involved in *V. cholerae*'s toxicity towards *D. discoideum* and macrophages (38, 80, 81). Moreover, this VgrG1-ACD was also responsible for intestinal inflammation and cholera toxin-independent fluid accumulation in an infant mouse model of infection (82). Furthermore, the VgrG1-ACD of the 7th pandemic *V. cholerae* strain C6706 was implicated in alternating intestinal peristalsis of zebrafish larvae, leading to the expulsion of preinoculated commensal bacteria (83). However, this effect was only observable upon constitutive T6SS expression using a genetically engineered derivative of this pandemic strain in which T6SS production occurred based on artificial expression of the gene encoding the quorum sensing- and chitin-linked transcription factor QstR (72, 73, 83–85).

Given this previous work on toxigenic strains, we tested whether the ACD of the clade C and D environmental isolates was likewise causative of the observed anti-amoebal response. To do so, we first generated truncated versions of VgrG1 that lacked the evolved ACD domain-containing C-terminus (Figure 4.2e). Importantly, these vgrG1ΔACD strains maintained their full anti-bacterial competitiveness (Figure 4.2f), which indicates that the ACD deletion did not impact the general assembly and/or activity of the T6SS machinery. However, as shown in Figure 4.2g, the amoebal grazing ability was restored on the lawns formed by the two vgrG1ΔACD environmental *V. cholerae* strains and resulted in equal (e.g., for strain SL6Y) or intermediate (e.g., strain SA10G) amoebal plaque numbers compared with those numbers observed for complete T6SS-defective mutants. The non-pandemic but toxigenic control strain ATCC25872 (Supplementary table 4.1; (29)), which, like strain V52, is constitutive T6SS active (64), also lost its anti-amoebal activity in the absence of the T6SS core structure or when the ACD of VgrG1 was missing (Figure 4.2g). We therefore concluded that some environmental *V. cholerae* might have evolved ACD-extended VgrG-

encoding genes, as the ACD's anti-eukaryotic activity protects them from environmental grazers. Whether pandemic strains subsequently horizontally acquired the ACD encoding region as previously suggested (86), or whether the last common ancestor between the pandemic and environmental lineages already contained this specialty that was later then lost from some wild strains, is currently unclear. We did observe, however, that the VgrG1-ACD protein is highly conserved between the pandemic strain A1552 and the environmental isolates, regardless of their clade (Supplementary table 4.4). Indeed, while the full length VgrG1-ACD protein showed >97% identity when compared among these strains, the ACD part alone showed identity levels above 99% with several strains sharing 100% identical ACD domains (Supplementary table 4.4).

4.4.6 Diversity of T6SS effectors in environmental *V. cholerae* isolates

As briefly mentioned in the previous section, the ACD-minus mutant of the clade D strain SA10G showed reduced but still significant residual T6SS-dependent anti-amoebal activity. We therefore wondered whether this could be explained by the presence of another anti-eukaryotic effector in one or several of the T6SS clusters. Furthermore, we were also interested in characterizing the full set of E/I modules in these strains, which would allow us to speculate about the strains' competitive potential against one another. Using the newly assembled genomic data together with the previously reported genome sequence of clade B strain SA5Y (69, 87) we determined the E/I modules of these 15 environmental strains and predicted their function based on BlastP analyses (Figure 4.3 and Supplementary table 4.3). Moreover, to arrange the E/I pairs into putative compatible groups, we defined their modules based on a percentage amino-acid identity of at least 30%, which is a typing approach previously applied (46, 86). Interestingly, we found evidence for several orphan immunity genes, meaning immunity genes that no longer coexisted with and were adjacent to a cognate effector-encoding gene, as had been previously reported by Kirchberger and

colleagues (86). These orphan genes were located in the 3' regions of the T6SS clusters and were likewise classified according to their module type (Figure 4.3 and Supplementary table 4.3). The characterization of the E/I modules and orphan immunity loci was restricted to the well-characterized T6SS clusters of *V. cholerae* (e.g., the large cluster as well as auxiliary clusters 1 and 2), as the reported auxiliary cluster 3 (E/I pair TseH/TseI; (58)) was absent from all environmental isolates. This finding is consistent with a preprinted study showing that this auxiliary cluster represents a pandemic strain-associated mobile genetic element (88) (Figure 4.3 and Supplementary table 4.3). Moreover, the recently characterized auxiliary clusters 4 and 5 were also absent from the environmental isolates (89, 90). Instead, automatic annotations followed by a manual inspection identified the presence of a novel T6SS cluster in strain SP7G (clade C), which we defined as auxiliary cluster 6 (Figure 4.3 and 4.4; Supplementary table 4.3). This cluster resides in the small chromosome 2 of strain SP7G and the gene order resembles that of auxiliary clusters 1 and 2, namely: *hcp*, *vgrG*, a putative adaptor-protein encoding gene (coding for a DUF4123 domain as described for Tap1 and VasW; (52, 53)), a gene of unknown function and a noncanonically ordered immunity/effector (I/E) pair, located 250 bp downstream of the previous gene (Figure 4.4a). BlastP analysis identified the predicted effector as a DUF2235-containing protein, similar to the effector protein encoded in auxiliary cluster 1. However, an alignment of the sequences of these two effector proteins (encoded in aux 1 or aux 6 of strain SP7G) showed only a 18% identity. Importantly, this cluster is located inside a larger genomic island. Our search in the PATRIC nucleotide sequence database (91) suggested that this aux 6 cluster (and the genomic island) is prominently represented in several *V. cholerae* strains, including 2013 environmental isolates from Bangladesh and an O35 strain (1311-69) isolated in 1969 from a patient in India (92). Our experimental investigation showed, that the auxiliary cluster 6 is active under laboratory conditions, as strain SP7G was able to kill its own kin when the latter

lacked the region that included the putative I/E gene pair (Figure 4.4b). Furthermore, SP7G lost some of its killing potential against *E. coli* when the I/E-including region of the auxiliary cluster 6 was missing (Figure 4.4c).

In the large T6SS cluster, pandemic *V. cholerae* strains, such as A1552 carry an A-type E/I module (Figure 4.3 and Supplementary table 4.3), for which, in fact, the peptidoglycan cell wall degradation effector corresponds to the C-terminal domain of the evolved VgrG3 protein (48, 51). When analyzing this locus in the environmental *V. cholerae* isolates, we noticed that the evolved nature of VgrG3 was conserved among these strains, except for clade C strain SP7G in which VgrG3 is solely a structural T6SS component that is followed by a B-type E/I module. How this effector is attached to the VgrG tip protein is, however, unclear, as no adaptor protein such as those encoded by *tap1* and *vasW* could be identified. This B-type effector is predicted to have a cellular adhesion function (46). Interestingly, strains W10G (clade A) and SA3G (clade B) contain a pandemic-like A-type E/I module (Figure 4.3), while the other strains carried a wide variety of E/I pairs, as described in detail in Supplementary table 4.3. Moreover, even though the amino acid identity of these effectors is below 30% when comparing different types, most of them have a common predicted function, namely, peptidoglycan degradation (Supplementary table 4.3). Therefore, with the exception of strain SP7G, all the environmental isolates have a dedicated anti-bacterial E/I module in the T6SS large cluster (Figure 4.3 and Supplementary table 4.3). Interestingly, pairwise comparisons of the effector and immunity protein sequences among different strains showed that strains harboring E/I modules from the same family had 100% immunity identity (Supplementary table 4.5). Given the diversity of E/I modules in this cluster when comparing all strains, the complete identity of immunity proteins from the same family could indicate recent acquisition by horizontal gene transfer.

Furthermore, L6G and SL6Y (clade C) are the only strains that also harbor one or several orphan immunity loci after the E/I module in this large T6SS cluster (Figure 4.3).

These orphan immunity loci are distinct in their type when compared with the current resident E/I modules. When we searched the PATRIC translated nucleotide sequence database (91) using the protein sequence from these orphan loci as the query, we noticed that these genes are only found as orphan loci in other genomes (i.e., they were located downstream of another E/I pair). The only exception was the second orphan locus in strain SL6Y, which encodes a G-type immunity protein (marked by "#" in Figure 4.3). The predicted protein showed 98.4% identity to true immunity proteins (e.g., those encoded directly adjacent to an effector gene) from several *V. cholerae* strains. Among those was strain 2633-78, an O1 CTX-negative isolate collected from sewage in Brazil in 1978. Interestingly, this strain was experimentally tested in a previous study, where it was shown to have an active T6SS (28).

The auxiliary cluster 1 contains the structural or evolved (e.g., encoding C-terminal ACD) *vgrG1* gene, as mentioned above (Figure 4.2 and 4.3). In addition, in the case of pandemic *V. cholerae*, this cluster harbors an A-type E/I module encoding the lipase effector TseL with anti-bacterial and anti-eukaryotic activity (47, 51, 54) followed by a C-type orphan immunity gene (86). Interestingly, all of the 15 environmental strains harbor C-type E/I modules as part of this auxiliary cluster 1, even though there is considerable polymorphism in the effector and immunity proteins (Figure 4.3 and Supplementary table 4.6). C-type effectors have a predicted alpha/beta hydrolase domain (DUF2235), which has been previously associated with T6SS phospholipases from *E. coli* and *Pseudomonas aeruginosa* (54, 89, 93). Furthermore, with the exception of strains W10G (clade A) and SO5Y (clade B), all environmental isolates also carry between two and five orphan immunity loci downstream of the E/I pair (Figure 4.3). Importantly, even though these orphan loci are C-

type as the *bona fide* E/I pair, they don't appear as duplications of the resident immunity gene, as they only share 60-85% identity with the latter. Considering the wide variety of E/I types in the large cluster as well as in auxiliary cluster 2 (see below), the apparent conservation of C-type E/I modules in the auxiliary cluster 1 of these strains, in addition to the many C-type orphan loci (as also observed in the *in silico* study by (86)) is quite intriguing and supports the idea that the C-type orphan immunity gene in the pandemic strains might be a remnant of a previous C-type E/I module of the strains' progenitor.

In pandemic *V. cholerae*, the auxiliary cluster 2 carries an A-type E/I module where the VasX effector is a pore-forming toxin (43, 55, 56). Our comparative genomic analyses showed that several of the environmental isolates likewise encode VasX-like effectors (A-type) at the same locus (Figure 4.3; Supplementary tables 4.3 and 4.7), while other strains carry D- and E-type E/I modules. While D-type effectors have been predicted to foster peptidoglycan degradation, E-type effectors are predicted to form pores, like VasX (46). We therefore reasoned that the residual ACD-independent T6SS-dependent anti-amoebal impact observed for strain SA10G (clade D) might be caused by this E-type effector from auxiliary cluster 2, especially as the latter is missing from clade C strain SL6Y in which removal of the ACD from VgrG1 was almost equivalent to a complete T6SS inactivation (Figure 4.2g). Interestingly, our analyses suggest that these E-type effectors also contain a common peptidoglycan-binding domain (e.g., an N-terminal Lysin Motif; (94)), which might render them bifunctional against bacteria and eukaryotes (Figure 4.3 and Supplementary table 4.3). Future work will address the exact characteristics of the putative E/I modules in more detail.

4.4.7 Competition among environmental strains occurs in the presence of nonidentical E/I modules

Previous studies showed that a plethora of T6SS-transported effectors have active anti-bacterial purposes and that cognate immunity proteins are required to protect the producer or its siblings from intoxication (43, 46, 47, 95). Strains with matching E/I modules could therefore coexist in the same environment (46), while competitive strains might clear a niche and propagate inside this niche in a clonal manner (96, 97). Importantly, as these E/I modules seem to move horizontally in an as yet uncharacterized manner (46, 98), the level of compatibility between strains will not follow their phylogenetic relatedness. We therefore sought to experimentally probe the strains' compatibility or competitiveness within this collection of environmental isolates. Indeed, while we scored T6SS activity against a laboratory strain of *E. coli* (see Figure 4.2), we wondered how the *V. cholerae* strains would behave when exposed to each other. Consistent with the fact that the E/I modules carried by the environmental strains are considerably different from those of pandemic *V. cholerae* (Figure 4.3 and Supplementary table 4.3), we experimentally demonstrated that these bacteria efficiently outcompeted the T6SS-silent pandemic strain A1552 (Figure 4.5a). Next, we tested interbacterial competition between the clade-representative strains in an assay in which the prey strains had their T6SS inactivated. As shown in Figure 4.5b, reciprocal killing of the T6SS-positive parental strains occurred and only kin strains were immune to the toxic assaults. However, as the E/I module types seemed more similar within clades than across clades, we extended our analyses and tested all pairwise combinations (this time, with both partners T6SS positive). The underlying rationale was that previous studies had identified different types of E/I modules *in silico* (46, 86, 89) but most pairwise competition assays were primarily linked to patient isolates and not to a larger collection of environmental isolates. Indeed, as described above and shown in Figure 4.3, our *in silico* prediction showed that strains within clades often contained similar E/I module types (Supplementary tables

4.5 to 4.7), even though the amino acid identity threshold for this categorization is rather low (30% as previously defined; (46, 86)).

As shown in Figure 4.5, we observed interesting and complex phenotypes in these extensive pairwise killing experiments. For instance, strains W6G and W7G from clade C were fully protected against each other's attacks (Figure 4.5 panels M and N). This finding is consistent with the 100% identity of all three T6SS clusters, including the orphan genes from aux cluster 1 (Supplementary tables 4.5 to 4.7), and their overall near clonality as described above. Clade D strains have the same E/I module types in all T6SS clusters. However, upon closer inspection of their E/I protein sequences, we can observe a 100% sequence identity among the three strains only in the large and aux 1 cluster-encoded proteins. For the aux 2 cluster, the encoded E/I proteins from strains E7G and SA7G are 100% identical, while the E/I pair carried by strain SA10G has only 95%/73% identity when compared with the other two strains' E/I pair (Supplementary tables 4.5 to 4.7). Consequently, and as shown in Figure 4.5 (panel O to Q), strains E7G and SA7G are fully compatible with each other, while strain SA10G can kill and be killed by the other two strains. We therefore concluded that the identity level in only one of the T6SS cluster-encoded immunity proteins causes the competitiveness among these strains.

Surprisingly, clade B strains SA5Y and SL4G have 100% identity in all three T6SS cluster-encoded immunity proteins (Supplementary tables 4.5 to 4.7) but are still able to kill each other with considerable efficiency (Figure 4.5, panels E and F). Why this is the case is currently unclear. However, we speculate that expression or immunity protein production might be impaired in those strains or that additional T6SS E/I modules are hidden in the strains' genomes, which were not easily identifiable based on a lack of the hallmark genes *paar*, *hcp* and *vgrG* in their vicinity.

Clade A strain W10G carries pandemic-like A-type E/I modules in the large and aux2 clusters (Figure 4.3). Interestingly, this strain can kill many of the other environmental strains remarkably well, such as clade B strains SA5Y, SL4G and SL5Y, and clade C strains SL6Y and SP6G. Notably, the T6SS active toxigenic strain V52 (which harbors A-type E/I modules in all T6SS clusters) was previously shown to outcompete strains carrying different combinations of E/I modules (46). Indeed, the AAA (A-type in all clusters) E/I modules is conserved in pandemic *V. cholerae*, even in those strains that caused former pandemics (e.g., 6th pandemic O1 classical strains), as well as in non-pandemic but toxigenic isolates. It was therefore speculated that this combination might be advantageous in a disease context (46). Clade B strain SA3G also contains A-type E/I modules in the large and aux 2 clusters, but it does not kill other strains at the same level as strain W10G. However, as noted above, this strain is even less efficient against laboratory *E. coli* prey strains (Figure 4.2A), which could mask its full effector toxicity potential.

Finally, a very interesting pairwise comparison is that of clade C strains SP7G and L6G. While L6G as a prey is very efficiently eliminated by strain SP7G (Figure 4.5, panel I), SP7G prey seemed almost resistant to L6G intoxication (Figure 4.5, panel L; please note, however, the commonly witnessed inter-experimental variation). Importantly, the observed phenotype was only mildly dependent on SP7G's own T6SS activity. As shown in Figure 4.4d, L6G is able to kill T6SS-deficient SP7G as a prey roughly 5-fold more efficiently compared to its T6SS-positive parental strain. These two strains contain E/I sets of different families in the large cluster and show 78% and 99.7% identity in the immunity proteins encoded in aux cluster 1 and 2, respectively. Based on these differences, we would expect that these strains fully compete with each other; however, killing almost exclusively occurs with SP7G as the predator. We therefore speculate that the large cluster-encoded K-type effector domains of strain L6G's evolved VgrG3 as well as SP7G's noncanonical structural

VgrG3 and its adjacently encoded putative effector protein might not be functional peptidoglycan destruction enzymes and therefore not necessarily active in interbacterial competition. Why a low level of protein identity in the aux cluster 2-encoded immunity protein is sufficient to protect strain SP7G from L6G assaults is currently unclear. It is also tempting to speculate that the *vice versa* interaction leads to L6G killing due to the additional auxiliary cluster 6 that is carried by strain SP7G. Upon experimental testing of this idea, we only observed a mild role for this additional auxiliary cluster with the SP7GΔAux6 strain showing slightly lower L6G killing activity compared to the SP7G WT parental strain (Figure 4.4e). Due to the overall inter-experimental variation, this difference was, however, not statistically significant. Further work is therefore necessary in order to delve deeper into these observed phenotypes.

Overall, closer inspection of all of the pairwise killing data attests to the complexity behind the T6SS compatibility code. As mentioned previously, even though some strains might harbor the same E/I families in the T6SS clusters, pairwise comparisons of these proteins shows that quite frequently their identity is not 100% (Supplementary tables 4.5 to 4.7), which appears to be necessary to allow coexistence (also observed in (97)). Our data therefore support what Unterweger and colleagues (46) initially speculated, namely, that the compatibility between strains seems to follow the level of polymorphism of their immunity proteins and that diversity in only one cluster-encoded protein is sufficient to drive competition.

4.5 CONCLUSION

In this study, we investigated two minor virulence factors, the pore forming hemolysin and the T6SS, in a set of fifteen environmental *V. cholerae* strains. We assessed the relevance

of these molecular weapons as defense mechanisms against amoebal predation (hemolysin and T6SS) and in the context of bacterial warfare (T6SS only). We showed that all of these environmental isolates possess a constitutively active T6SS and are able to use the machinery as a bacterial killing device. In contrast, only a subset of these strains was able to efficiently suppress grazing by *D. discoideum* amoebae, a phenotype that was dependent on the eukaryote-specific ACD of the evolved VgrG1 T6SS effector. Careful *in silico* identification unveiled an extensive T6SS repertoire of E/I pairs and orphan immunity loci. Consistent with this finding, we observed extensive interbacterial competition under pairwise coculture conditions whereby mutual compatibility was rarely achieved. Importantly, our study also confirmed that both molecular weapons, the T6SS and the hemolysin toxin, are constitutively active in the environmental isolates. It is expected that these *V. cholerae* strains constantly compete with other microorganisms for nutrients and space within the marine environment besides being under persistent predation pressure from bacterivorous grazers. It seems therefore reasonable to assume that keeping their molecular weapons constitutively active might provide them with a competitive advantage in this natural habitat. Importantly, this is in sharp contrast to the tight regulation of these machineries in the well-studied pandemic patient isolates. Future work is therefore required to decipher how this differential production pattern is achieved in pandemic versus non-pandemic strains and whether this tight regulatory control might provide specific benefits to the former strains.

4.6 EXPERIMENTAL PROCEDURES

4.6.1 Bacterial strains and growth conditions

The bacterial strains (*V. cholerae*, *E. coli* and *K. pneumoniae*) used in this study are listed in Supplementary table 4.1. Unless otherwise stated, all strains were grown aerobically in

Lysogeny broth (LB; 10 g/L of tryptone, 5 g/L of yeast extract, 10 g/L of sodium chloride; Carl Roth) or on LB agar plates at 30°C. Half-concentrated defined artificial seawater medium (0.5×DASW) containing HEPES and vitamins (99) was used for growth on chitinous surfaces for strain construction based on chitin-induced natural transformation (see below).

D. discoideum amoebae (strain Ax2 Ka) were cultured in HL5 medium supplemented with glucose (Formedium, UK). For amoebal grazing assays (e.g., plaque formation assays; see below), SM/5 medium (final concentrations: 2 g/L of glucose, 2 g/L of bacto peptone, 2 g/L of yeast extract, 0.2 g/L of $MgSO_4$ $7H_2O$, 1.9 g/L of KH_2PO_4, 1 g/L of K_2HPO_4; pH 6.4), was mixed with 2% agar to prepare SM/5 plates (20 mL/plate). $CaCl_2$ (50 µM)-supplemented Sörensen's buffer (8 g /4 L of KH_2PO_4, 1.16 g /4 L of Na_2HPO_4 ; pH 6; (100)) was used as the washing and resuspension buffer for the amoebae and the bacteria that were used in the plaque formation assay.

The following antibiotics were added if required at the given concentration: kanamycin (75 µg/ml), rifampicin (100 µg/ml), streptomycin (100 µg/ml) and chloramphenicol (2.5 µg/ml).

4.6.2 Genetic engineering

V. cholerae strains were genetically modified using chitin-induced transformation as previously described (101–103). This method relies on natural transformation triggered by growth on chitin followed by the addition of a PCR fragment that carried the desired genetic change. To achieve higher numbers of transformants, the protocol was slightly modified. The PCR fragments were added twice (24h and 36h after bacterial inoculation on the chitin flakes) and cells were enriched in 2×YT medium (Carl Roth) before selective plating. PCR amplifications were conducted using Pwo (Roche) and GoTaq (Promega) polymerases

according to the suppliers' recommendations. Following initial screening by PCR (using bacterial cells as the templates), genetically engineered loci were verified by Sanger sequencing (Microsynth, Switzerland).

The rifampicin-sensitive *V. cholerae* strain A1552-Rif^S was generated by a combination of natural cotransformation (104) and our previously described counter selectable Trans2 approach (64). To this end, a 4kb PCR fragment was amplified, harboring a mutation, which restored the native *rpoB*-encoded protein (F531S substitution) from its mutated version in the parental strain A1552 (RpoB[S531F]; (69)). *V. cholerae* A1552 was then cotransformed with this fragment and another 3,924 bp fragment containing flanking regions matching *lacZ* and two selection markers (*aph* and *pheS**; Supplementary table 4.1). Transformants were selected on kanamycin-containing agar plates and the lost rifampicin resistance was scored based on replica plating on plates +/- rifampicin using a velvet cloth. A second round of natural transformation followed to restore the *lacZ* gene by adding a WT *lacZ* PCR fragment to chitin-grown cells followed by a counter selection of the *pheS** allele on 4-chloro-phenylalanine (20mM)-containing agar plates as previously described (64). To confirm the restoration of native *rpoB*, genomic DNA of A1552 Rif^S was isolated and the PCR-amplified *rpoB* gene was Sanger-sequenced.

4.6.3 Amoebal grazing assay

To determine the predatory capacity of *D. discoideum* on bacterial lawns of *V. cholerae*, plaque formation was scored following a previously described protocol (38) with minor modifications. Briefly, bacteria were cultured overnight in LB medium at 30°C and harvested by centrifugation. The cell pellet was washed and resuspended in SorC buffer (100), and then diluted with SorC to reach an optical density at 600 nm (OD_{600}) of 5.5 in a final volume

of 400 µL. Cultured *D. discoideum* amoebae (in HL5 with glucose; Formedium, UK) were detached from culture dishes using cell scraper (SPL Life Sciences) and collected by centrifugation (3 min 1,000 rcf), resuspended in SorC buffer, and enumerated in a KOVA counting chamber (KOVA International, USA). The amoebal concentration was adjusted to 2×10^4 cells/ mL and 20 µL of this suspension (corresponding to ~400 amoebal cells) was mixed with the 400 µL of bacterial suspension. The mixture was gently spread on two parallel SM/5 plates using a plastic rake (VWR), resulting in technical replicates. The plates were wrapped in aluminum foil and incubated at 24°C for 5 days. After this incubation period, *D. discoideum* plaque numbers were enumerated. As a positive control, we included a frequently used nonencapsulated *Klebsiella* strain (105) for which the resulting plaque numbers were set to 100%. Three biologically independent experiments were performed. The individual experimental data points (mean of technical replicates) as well as the overall average of the independent experiments (+/- standard deviation) are shown in each graph. A two-tailed Student's *t*-test was performed to determine statistical significance.

4.6.4 Hemolysin activity

The hemolytic activity of *V. cholerae* was assayed using trypticase soy agar containing 5% sheep blood (BD, Heidelberg, Germany). To do so, the respective overnight cultures were spotted (2 µL) onto the plates and incubated at 30°C for 24 h, after which pictures of the plates were taken.

4.6.5 Interbacterial killing assays

Bacterial killing was assessed following a previously established assay with minor modifications (72). The prey cells (*E. coli* or *V. cholerae*, as indicated) and the respective

predator bacteria were mixed at a ratio of 1:10 and spotted onto paper filters on prewarmed LB agar plates. After 4 h of incubation at 37°C, the bacteria were resuspended, serially diluted, and spotted onto antibiotic-containing (rifampicin or streptomycin) LB agar plates to enumerate the colony-forming units (shown as CFU/ml). The majority of these killing experiments were performed using exponentially growing *V. cholerae* (OD_{600} ~1). For the pairwise killing experiments of all environmental strains, OD_{600}-adjusted overnight cultures were used, which resulted in biologically similar outcomes to the samples derived from exponentially growing cultures. Statistically significant differences were determined on log-transformed data (106) by a two-tailed Student's *t*-test of three biologically independent replicates. If no prey bacteria were recovered, the value was set to the detection limit to allow for statistical analysis.

4.6.6 SDS-PAGE and western blotting

To check the production of the Hcp protein, cell lysates were prepared as described previously (76). In brief, exponentially growing bacteria (~3 hours of growth after a 1:100 back dilution from overnight cultures) were pelleted and then resuspended in Laemmli buffer, adjusting for the total number of bacteria according to the cultures' OD_{600} values. To check for T6SS-secreted Hcp, 1.5 ml of the culture supernatant was filter sterilized (0.2-μm filter; VWR) and the proteins were precipitated using trichloroacetic acid (TCA). The precipitated proteins were washed with acetone before being resuspended in 30 μL of Laemmli buffer. All samples were heated at 95°C for 15 min.

Proteins were separated by sodium dodecyl sulfate (SDS)-polyacrylamide gel electrophoresis (PAGE) using 15% gels and then western blotted as previously described (107). Primary antibodies against Hcp (Eurogentec; (76)) were used at 1:5,000 dilution while

anti-Sigma70-HRP antibodies (BioLegend, USA distributed via Brunschwig, Switzerland) were diluted 1:10,000 and served as a loading control. Goat anti-rabbit horseradish peroxidase (HRP) (diluted at 1:20,000; Sigma-Aldrich, Switzerland) was used as the secondary antibody against the anti-Hcp primary antibody. Lumi-Light^{PLUS} western blotting substrate (Roche, Switzerland) served as the HRP substrate. The signals were detected using a ChemiDoc XRS+ station (BioRad).

4.6.7 Preparation of genomic DNA for whole-genome sequencing

Genomic DNA (gDNA) was purified from 2 ml of an overnight culture of the respective strain. DNA extraction was performed using 100/G Genomic-tips together with a Genomic DNA buffer set as described in the manufacturer's instructions (Qiagen). After precipitation, the DNA samples were washed twice with cold 70% ethanol and dissolved in Tris buffer (10 mM Tris-HCl, pH 8.0).

4.6.8 Long-read PacBio genome sequencing

Sample preparation and genome sequencing was performed by the Genomic Technology Facility of the University of Lausanne (Switzerland) using standard protocols. Briefly, DNA samples were sheared in Covaris g-TUBEs to obtain fragments with a mean length of 20 kb. The sheared DNA was used to prepare each library with the PacBio SMRTbell template prep kit 1 (Pacific Biosciences) according to the manufacturer's recommendations. The resulting library was size selected on a BluePippin system (Sage Science, Inc.) for molecules larger than 15 kb, which excluded smaller plasmids. Each library was sequenced on one single-molecule real-time (SMRT) cell with P6/C4 chemistry and MagBeads on a PacBio RS II system at a movie length of 360 min. Genome assembly was performed using

the protocol RS_HGAP_Assembly.3 in SMRT Pipe 2.3.0, and circularization of the genomes was achieved using the Minimus assembler of the AMOS software package 3.1.0 using default parameters (108). The assembled genomes were initially annotated using Prokka 1.12 (109) but due to several incompatibilities with the NCBI database, they were reannotated with their own pipeline (PGAP annotation) during NCBI submission. The genomic data and NCBI accession numbers are summarized in Supplementary table 4.2.

Notably, we observed minor differences in the T6SS clusters of the strains when comparing Prokka and PGAP annotations. For instance, we noticed slight differences in the starting points of some of the predicted orphan loci from aux 1, even though the reading frames were comparable between the annotation methods. Additionally, the *vgrG3* gene from the large cluster of some strains was shortened in the PGAP annotation, which most likely excluded the effector encoding part of the gene. In these cases (indicated in Supplementary table 4.3), we considered the Prokka annotation for our analysis. Finally, some annotations were created beyond internal frameshifted stop codons. For instance, in the auxiliary cluster 1 of strain W10G, the effector gene (HPY12_07370) contains several predicted stop codons in the PGAP annotation, due to the insertion of a C in position 323. A similar case was observed in the putative effector gene (HPY12_14010) from auxiliary cluster 2, due to the insertion of a C in position 889. Both of these regions were Sanger sequenced after PCR-amplification using the same genomic DNA samples as templates that were initially used for PacBio library preparation, which confirmed that the additional C bases were in both cases a sequencing artifact and the genes were properly maintained in strain W10G (as indicated in Figure 4.3).

4.6.9 Characterization of E/I modules from environmental strains

T6SS clusters of the environmental strains were identified by searching the Prokka-annotated genomes for conserved genes, such as *paar, vgrG* and *hcp,* as well as according to their location when aligned to the genome of the pandemic strain A1552 (69). All identified putative effectors were compared by BLAST against the NCBI database to identify conserved domains. Additional characterization was made based on the previous literature (46, 86). Furthermore, all E/I modules and orphan immunity loci were classified for their family type. Proteins with a sequence of less than 30% identity were considered as distinct incompatible types, as previously described (46, 86). For VgrG3 proteins, the typing was only based on the effector portion of the protein (46). To determine this part of the protein, the full VgrG3 sequence from all strains was aligned. The conserved region corresponding to the VgrG part of the protein was subsequently removed, which left only the variable C-terminal effector domain. This part was then used for typing and pairwise comparisons. The putative orphan immunity proteins were queried against a translated nucleotide database (PATRIC; (91)) to identify homologous *bona fide* immunity proteins encoded adjacent to an effector gene in the other T6SS clusters.

4.7 DATA AVAILABILITY

PacBio raw reads of the 14 whole-genome sequenced strains have been deposited in NCBI's Sequence Read Archive (SRA) under Bioproject accession number PRJNA633476. Details on the SRA accession numbers, BioSamples, and individual accession numbers of the *de novo* assembled and circularized genomes are provided in Supplementary table 4.2.

4.8 ACKNOWLEDGEMENTS

The authors acknowledge Sandrine Stutzmann for technical assistance, A. Boehm for sharing the published environmental *V. cholerae* isolates, Lisa Metzger for provision of strain constructs, and Thierry Soldati for sharing the *D. discoideum* amoebae. We also thank Frédérique Le Roux and current and former members of the Blokesch group and of the Lausanne/Geneva amoeba club for fruitful discussions. The authors acknowledge the staff of the Lausanne Genomic Technologies Facility at the University of Lausanne for sample processing, sequencing, and genome assembly, the NCBI submission staff for help with data deposition, and Anne-Catherine Portmann and Ivan Mateus for preliminary analysis of the genome data. We also note that the purchase of the Pacific Biosciences RS II instrument at the University of Lausanne was partially supported by the Loterie Romande through the Fondation pour la Recherche en Médecine Génétique. This work was supported by the Swiss National Science Foundation (310030_185022), the Novartis Foundation for medical-biological Research (#18C178), and a consolidator grant by the European Research Council (724630). MB is a Howard Hughes Medical Institute (HHMI) International Research Scholar (grant 55008726).

4.9 FIGURES

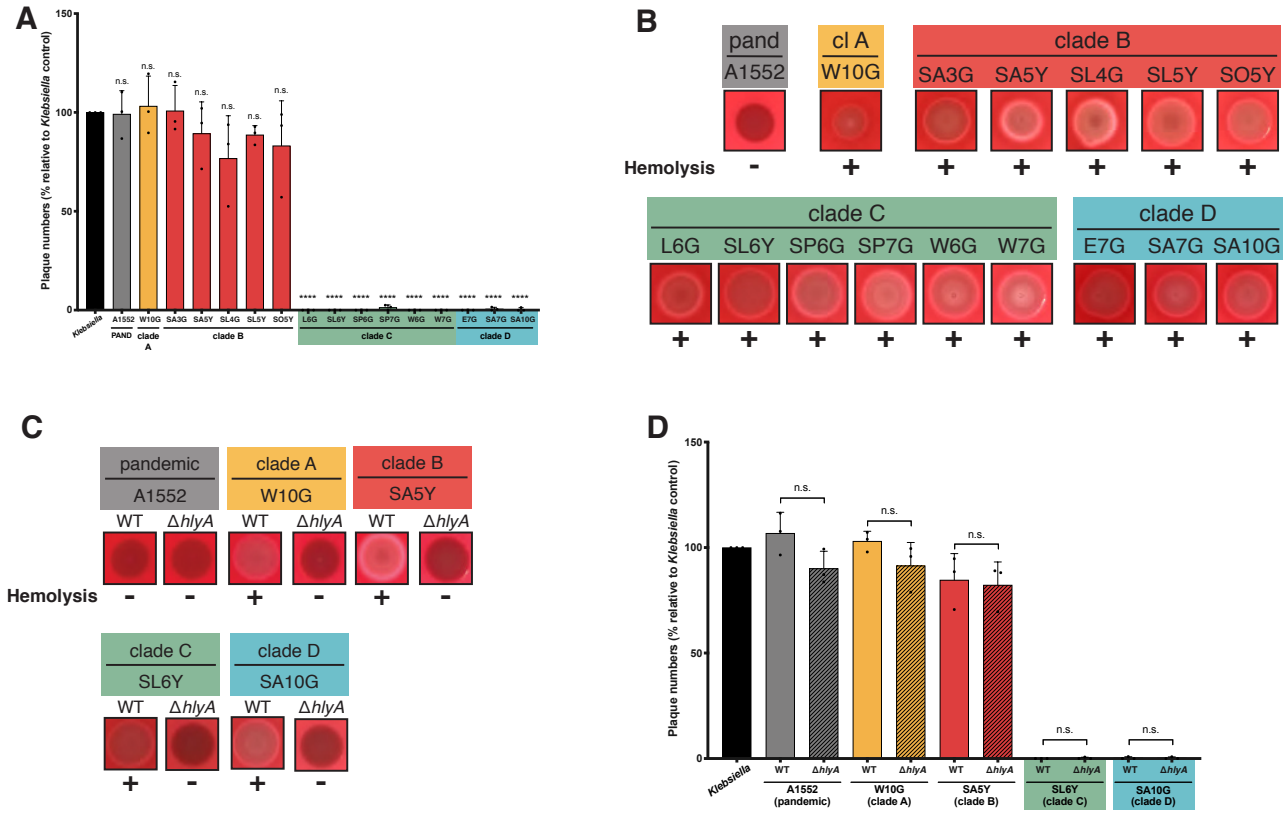


Figure 4.1 Defense against amoebal predation by a subset of *V. cholerae* strains.

(A, D) Amoebal predation was scored using *D. discoideum* grazing assays in which formed plaques on bacterial lawns were enumerated. Plaque numbers are indicated relative to those formed on a lawn of *K. pneumoniae*, which served as a positive control. Bar plots represent the average of at least three independent biological replicates (±SD). Statistical significance is indicated (n.s., not significant; **** $p < 0.0001$). **(B and C)** Hemolytic activity was tested on blood agar plates. Pandemic *V. cholerae* strain A1552 as well as all environmental isolates **(B)** or a representative subset together with their respective *hlyA*-minus derivatives **(C)** were assessed for hemolysis.

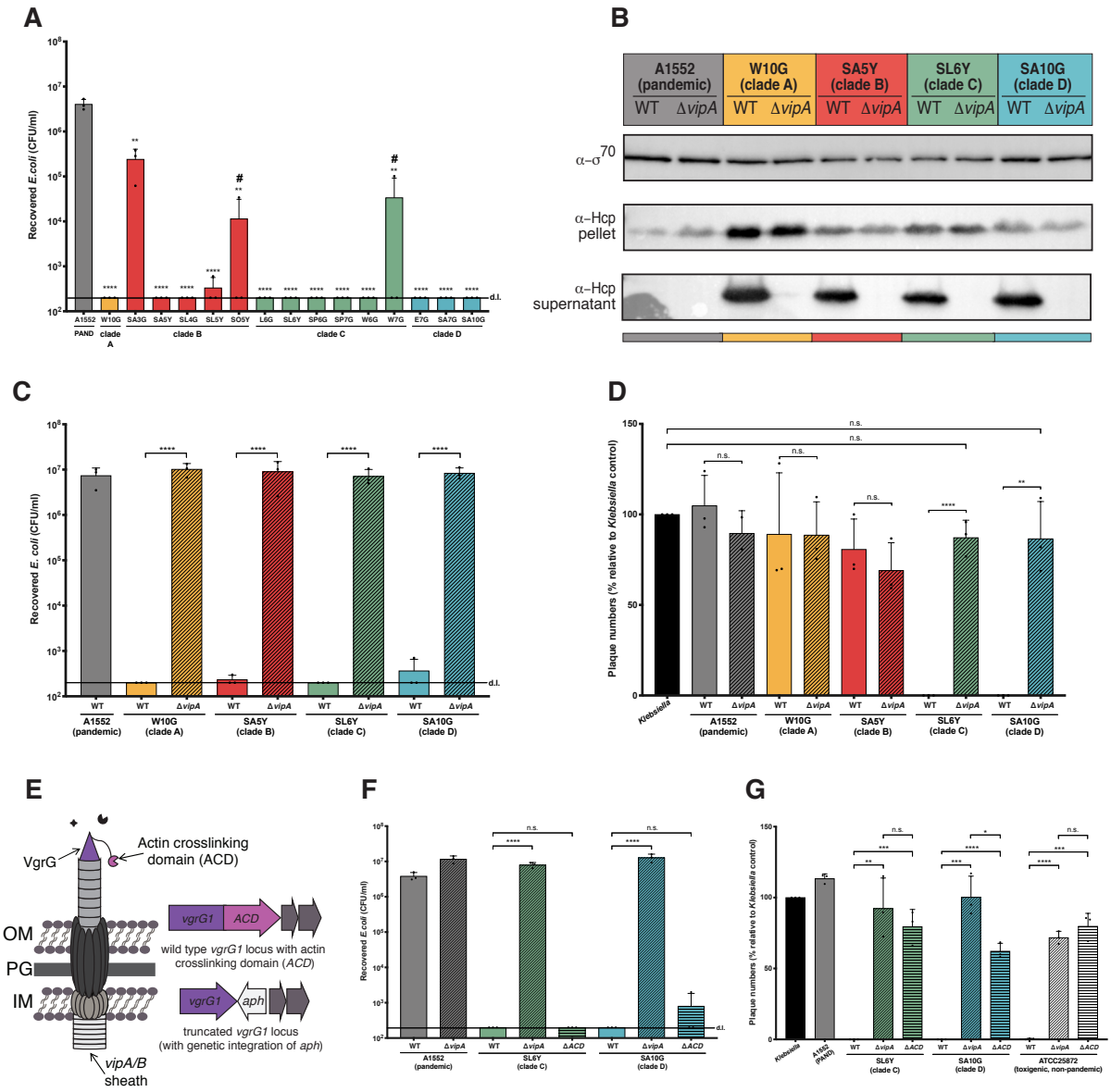


Figure 4.2 Constitutive T6SS activity linked to an ACD-containing effector inhibits amoebal grazing.

(A, C, F) Bacterial killing assays using *E. coli* as prey. Numbers of surviving prey are depicted on the Y-axis (CFU/ml). Statistical significance in panel **A** is shown above each strains' bar and calculated relative to the T6SS-silent negative control strain A1552. #, for these strains, the killing activity was only reduced in one of the three independent experiments. **(B)** T6SS activity in representative environmental strains. Hcp detection in WT and Δ*vipA* mutants of representative environmental isolates. Intracellular (pellet) and secreted (supernatant) Hcp were assessed by immunoblotting using Hcp-directed antibodies. Detection of σ70 served as a loading control. **(D, G)** T6SS- and ACD-dependency of the anti-amoebal defense. Plaque formation by *D. discoideum* on bacterial lawns formed by representative *V. cholerae* WT, *vipA* derivatives **(D and G)** and ACD-minus **(G)** strains. Details as in Fig. 1. The toxigenic non-pandemic strain ATCC25872 and its site-directed mutant served as control in panel **G**. **(E)** Simplified scheme of the T6SS. The actin crosslinking domain (ACD) consists of a C-terminal extension of the VgrG1 tip protein and this multidomain protein is encoded by the *vgrG1* locus (shown on the right). Removal of the ACD-encoding sequence was accomplished through site-directed integration of a stop codon concomitantly with an *aph* selective marker. Bar plots in all panels represent the average of at least three independent biological replicates (±SD). < dl, below detection limit. Statistical significance is indicated (n.s., not significant; * $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$; **** $p < 0.0001$; for panel A, each sample was compared to the A1552 control).

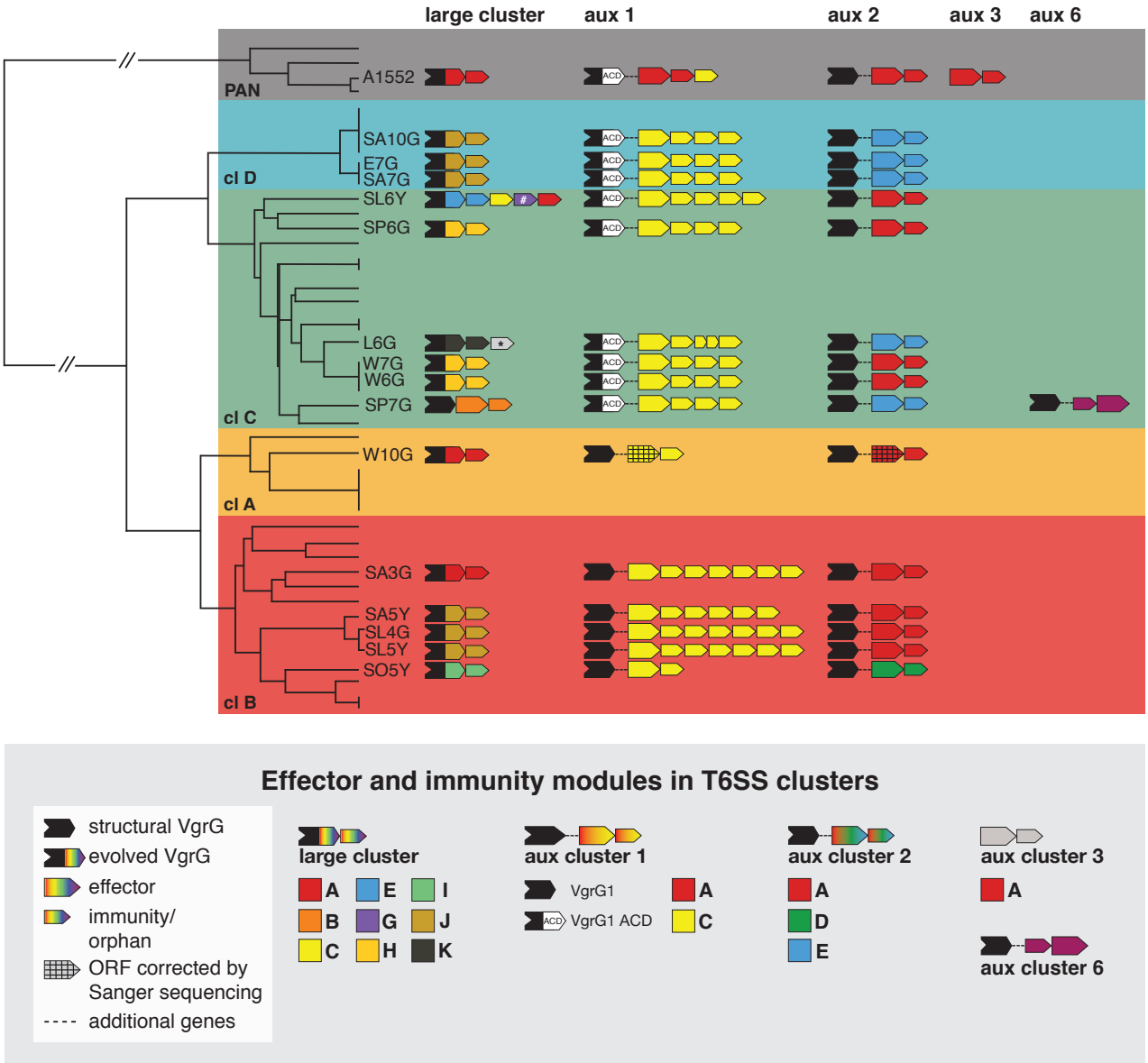


Figure 4.3 T6SS effector/immunity typing scheme of pandemic and environmental *V. cholerae*.

The previously published phylogenetic tree was freely adapted from (65) (not fully to scale). The clade color code is depicted in the background. The E/I type of each T6SS cluster (large cluster as well as auxiliary (aux) clusters 1, 2, 3 and 6) are schematized for each strain. Large black arrows symbolize *vgrG* genes, which were classified as structural (black) or evolved (colored tips, to represent the different types of C-terminal effector domains). Large and small colored arrows represent effector and immunity genes, respectively, according to the color code indicated in the legend below the scheme. Immunity genes that are not adjacent to a putative effector gene are considered as orphan immunity loci. * depicts an orphan immunity gene from strain L6G whose gene product was slightly below the amino acid identity threshold of 30% (20.3%) relative to C-type immunity proteins. # depicts an orphan immunity gene that is a homolog to *bona fide* immunity genes in other genomes. Effector genes in auxiliary clusters 1 and 2 of strain W10G (marked in figure with a pattern) were wrongly annotated in the PacBio genome sequence due to a frameshift sequencing artifact splitting each gene into two ORFs; this sequencing error was corrected by Sanger sequencing and the corrected single gene is shown in this figure.

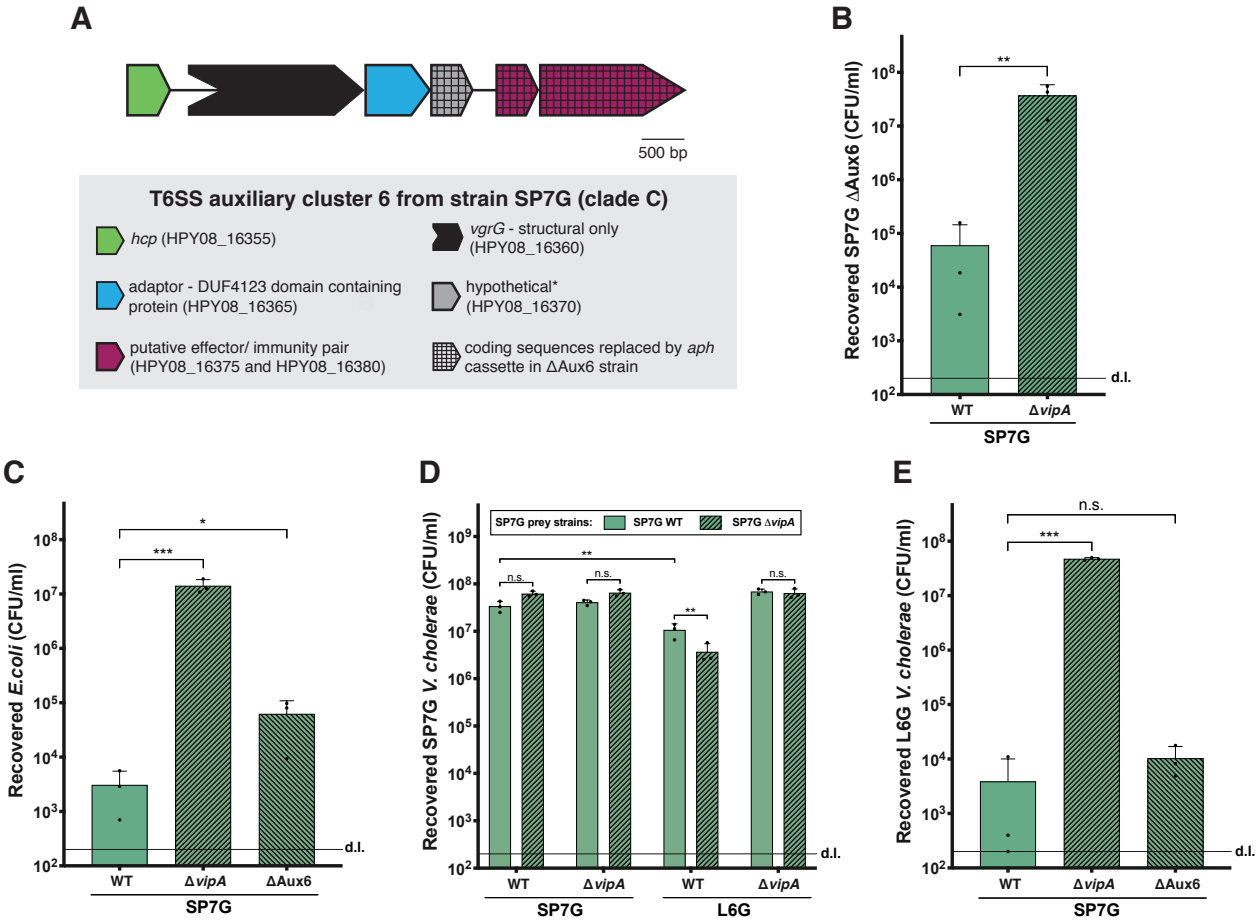


Figure 4.4 The T6SS auxiliary cluster 6 of strain SP7G is active.

(A) The 6.4-kb long auxiliary T6SS cluster 6 carried on the chromosome 2 of strain SP7G (locus tags HPY08_16355 to HPY08_16380 in CP053809) is depicted. Gene symbols are explained below the scheme. *Gene HPY08_16370 was annotated (PGAP method from NCBI) beyond internal stop codons, which were verified by Sanger sequencing (the initial Prokka annotation predicted two separate ORFs). The region spanning HPY08_16370 to HPY08_16380 was replaced by an *aph* cassette to test for aux 6 activity. **(B-E)** Bacterial killing assays using SP7GΔAux6 **(B)**, *E. coli* **(C)**, SP7G WT and Δ*vipA* **(D)**, or L6G **(E)** as prey. Numbers of surviving prey are depicted on the Y-axis (CFU/ml). Bar plots represent the average of three independent biological replicates (±SD). Statistical significance is indicated (n.s., not significant; * $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$).

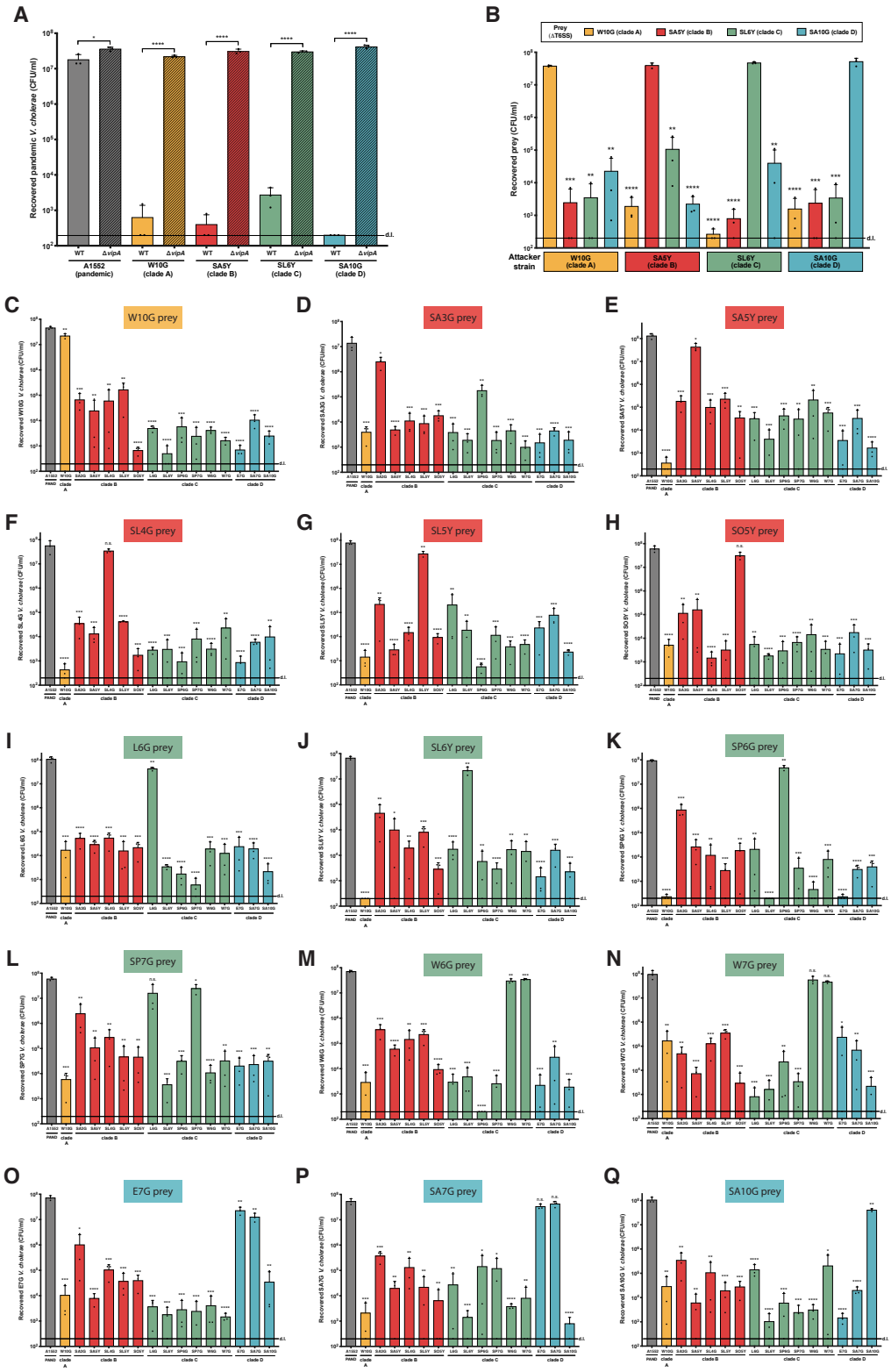


Figure 4.5 Competitiveness among environmental strains.

Graphs in all panels represent bacterial killing assays as described in Fig. 2 using the predator strains as indicated on the X-axes and as prey: **(A)** the pandemic *V. cholerae* strain A1552; **(B)** representative *vipA*-minus (ΔT6SS) mutants of the environmental isolates as shown above the graph; or **(C to Q)** each environmental isolate as shown in the graph title for each panel and on the Y-axis. Plots represent the average of three independent biological replicates (±SD). < dl, below the detection limit. Statistical significance is indicated (n.s., not significant; * $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$; **** $p < 0.0001$).

4.10 SUPPLEMENTARY TABLES

Supplementary table 4.1 *Vibrio cholerae, Escherichia coli* and *Klebsiella pneumoniae* strains used in this study.

Strain names	Genotype / description*	Internal strain number	Reference
V. cholerae strains			
A1552 (WT)	Wild-type, O1 El Tor Inaba; Rif^R	MB_1	(68)
W10G	Environmental isolate (clade A) collected in Waddell Creek (CA, USA) in October 2004	MB_5537	(65)
SA3G	Environmental isolate (clade B) collected in Old Salinas River (CA, USA) in March 2004	MB_957	(65)
SA5Y	Environmental isolate (clade B) collected in Old Salinas River (CA, USA) in May 2004	MB_353	(65)
SL4G	Environmental isolate (clade B) collected in San Lorenzo River (CA, USA) in April 2004	MB_955	(65)
SL5Y	Environmental isolate (clade B) collected in San Lorenzo River (CA, USA) in May 2004	MB_954	(65)
SO5Y	Environmental isolate (clade B) collected in Soquel Creek (CA, USA) in May 2004	MB_960	(65)
L6G	Environmental isolate (clade C) collected in Lagunitas Creek (CA, USA) in June 2004	MB_956	(65)
SL6Y	Environmental isolate (clade C) collected in San Lorenzo River (CA, USA) in June 2004	MB_953	(65)
SP6G	Environmental isolate (clade C) collected in San Pedro Creek (CA, USA) in June 2004	MB_964	(65)
SP7G	Environmental isolate (clade C) collected in San Pedro Creek (CA, USA) in July 2004	MB_952	(65)
W6G	Environmental isolate (clade C) collected in Waddell Creek (CA, USA) in June 2004	MB_354	(65)
W7G	Environmental isolate (clade C) collected in Waddell Creek (CA, USA) in July 2004	MB_962	(65)
E7G	Environmental isolate (clade D) collected in Moss Landing Harbor (CA, USA) in July 2004	MB_963	(65)
SA7G	Environmental isolate (clade D) collected in Old Salinas River (CA, USA) in July 2004	MB_959	(65)
SA10G	Environmental isolate (clade D) collected in Old Salinas River (CA, USA) in October 2004	MB_5539	(65)
A1552Δ*hlyA*	A1552 with *aph* cassette in *hlyA* (A1552VC_A02993; homolog to VCA0219*); Rif^R, Kan^R	MB_3934	This study
W10GΔ*hlyA*	W10G with *aph* cassette in *hlyA* (HPY12_14900); Kan^R	MB_8548	This study
SA5YΔ*hlyA*	SA5Y with *aph* cassette in *hlyA* (VC-Sa5Y_02920); Kan^R	MB_4622	This study
SL6YΔ*hlyA*	SL6Y with *cat* cassette in *hlyA* (HPY06_14140); Cm^R	MB_8549	This study
SA10GΔ*hlyA*	SA10G with *aph* cassette in *hlyA* (HPY11_14585); Kan^R	MB_8550	This study
A1552Δ*vipA*	A1552 with *aph* cassette in *vipA* (A1552VC_A02892; homolog to VCA0107); Rif^R, Kan^R	MB_3012	This study
W10GΔ*vipA*	W10G with *aph* cassette in *vipA* (HPY12_14395); Kan^R	MB_6850	This study
SA5YΔ*vipA*	SA5Y with *aph* cassette in *vipA* (VC-Sa5Y_02822); Kan^R	MB_5772	This study
L6GΔ*vipA*	L6G with *aph* cassette in *vipA* (HPY16_13230); Kan^R	MB_8628	This study
SL6YΔ*vipA*	SL6Y with *aph* cassette in *vipA* (HPY06_13635); Kan^R	MB_8551	This study
SP7GΔ*vipA*	SP7G with *aph* cassette in *vipA* (HPY08_13415); Kan^R	MB_8629	This study
SP7GΔ*vipA*-Rif^R	SP7G with *aph* cassette in *vipA* (HPY08_13415); transformed with PCR of *rpoB* gene containing mutation (S531F) that confers rifampicin resistance; Kan^R, Rif^R	MB_8630	This study
SA10GΔ*vipA*	SA10G with *aph* cassette in *vipA* gene (HPY11_14100); Kan^R	MB_8552	This study

A1552-Rif^{S}	A1552 sensitive to Rifampicin (please check Material and Methods for details on construction)	MB_5161	This study
A1552Δ*lacZ*::FRT -Kan-pheS*-FRT	A1552 with construct Δ*lacZ*::FRT-Kan-pheS*-FRT; Rif^{R}, cPhe^{S}	MB_4609	(64)
A1552-Rif^{S}Δ*vipA*	A1552-Rif^{S} with *aph* cassette in *vipA*; Rif^{R}, Kan^{R}	MB_8553	This study
SL6Y-vgrG1ΔACD	SL6Y with truncated *vgrG1* through insertion of a stop codon and insertion of an *aph* cassette that replaces the ACD-encoding sequence; Kan^{R}	MB_8557	This study
SA10G-vgrG1ΔACD	SA10G with truncated *vgrG1* through insertion of a stop codon and insertion of an *aph* cassette that replaces the ACD-encoding sequence; Kan^{R}	MB_8558	This study
ATCC25872	*V. cholerae* non-O1 strain (O37); isolated in 1965, Czechoslovakia; intermediate resistant to Strep	MB_276	(29)
ATCC25872Δ*vipA*	ATCC25872 deleted for *vipA* (TransFLP); intermediate resistant to Strep	MB_5065	(64)
ATCC25872 -vgrG1ΔACD	ATCC25872 with truncated *vgrG1* through insertion of a stop codon and insertion of an *aph* cassette that replaces the ACD-encoding sequence; Kan^{R}, intermediate resistant to Strep	MB_8559	This study
W10G-Rif^{R}	W10G selected for rifampicin resistance	MB_8562	This study
SA3G-Rif^{R}	SA3G transformed with PCR of *rpoB* gene containing mutation (S531F) that confers rifampicin resistance	MB_8563	This study
SA5Y-Rif^{R}	SA5Y selected for rifampicin resistance	MB_8564	This study
SL4G-Rif^{R}	SL4G selected for rifampicin resistance	MB_8565	This study
SL5Y-Rif^{R}	SL5Y selected for rifampicin resistance	MB_8566	This study
SO5Y-Rif^{R}	SO5Y selected for rifampicin resistance	MB_8567	This study
L6G-Rif^{R}	L6G transformed with PCR of *rpoB* gene containing mutation (S531F) that confers rifampicin resistance	MB_8568	This study
SL6Y-Rif^{R}	SL6Y selected for rifampicin resistance	MB_8569	This study
SP6G-Rif^{R}	SP6G selected for rifampicin resistance	MB_8570	This study
SP7G-Rif^{R}	SP7G transformed with PCR of *rpoB* gene containing mutation (S531F) that confers rifampicin resistance	MB_8571	This study
W6G-Rif^{R}	W6G selected for rifampicin resistance	MB_8572	This study
W7G-Rif^{R}	W7G selected for rifampicin resistance	MB_8573	This study
E7G-Rif^{R}	E7G selected for rifampicin resistance	MB_8574	This study
SA7G-Rif^{R}	SA7G selected for rifampicin resistance	MB_8575	This study
SA10G-Rif^{R}	SA10G selected for rifampicin resistance	MB_8576	This study
SP7GΔAux6	SP7G with *aph* cassette in the end of T6SS Aux 6 (HPY08_16370-16380); Kan^{R}	MB_8631	This study
SP7GΔAux6-Rif^{R}	SP7G with *aph* cassette in the end of T6SS Aux 6 (HPY08_16370-16380); Kan^{R}; Rif^{R}	MB_8632	This study
***E. coli* and *K. pneumoniae* strains**			
TOP 10	F- mcrA Δ(mrr-hsdRMS-mcrBC) φ80lacZΔM15 ΔlacX74 nupG recA1 araD139 Δ(ara-leu)7697 galE15 galK16 rpsL(StrR) endA1 λ-	MB_741	This study
K. pneumoniae	*Klebsiella* sp. / pANT5 (GFP)	MB_6380	(105)

*reference locus tags belong to reference strain N16961 according to (67).

Supplementary table 4.2 Information of the long-read whole genome sequencing data and assemblies of NCBI BioProject PRJNA633476.

Strain name	Strain ID	BioSample	GenBank accession numbers	Number of bases	Number of reads	Mean read length	Total number of contigs	Contig length after circularization	Total genome size	Mean coverage	GC% content
W10G	MB#5537	SAMN14944057	CP053794 (chr1) CP053795 (chr2)	2,050,336,308 bp	119,162	17,206 bp	2 (chr1+chr2)	3,033,188 bp (chr1) 1,028,867 bp (chr2)	4,062,055 bp	292 x	47.7% (chr1) 47.2% (chr2)
SA3G	MB#957	SAMN14944058	CP053744 (chr1) CP053745 (chr2)	1,239,715,817 bp	70,375	17,615 bp	2 (chr1+chr2)	3,020,084 bp (chr1) 1,159,698 bp (chr2)	4,179,782 bp	201 x	47.6% (chr1) 47.0% (chr2)
SA5Y#	MB#353	SAMN08813238	CP028892 (chr1) CP028893 (chr2)	635,540,812 bp	35,390	17,958 bp	2 (chr1+chr2)	2,955,400 bp (chr1) 1,095,478 bp (chr2)	4,050,878 bp	133 x	47.8% (chr1) 46.8% (chr2)
SL4G	MB#955	SAMN14944065	CP053796 (chr1) CP053797 (chr2)	1,083,962,388 bp	55,665	19,472 bp	2 (chr1+chr2)	2,954,462 bp (chr1) 1,173,143 bp (chr2)	4,127,605 bp	186 x	47.8% (chr1) 46.4% (chr2)
SL5Y	MB#954	SAMN14944067	CP053798 (chr1) CP053799 (chr2)	1,623,840,987 bp	84,121	19,303 bp	2 (chr1+chr2)	2,947,299 bp (chr1) 1,086,276 bp (chr2)	4,033,575 bp	336 x	47.8% (chr1) 46.9% (chr2)
SO5Y	MB#960	SAMN14944069	CP053800 (chr1) CP053801 (chr2)	865,600,319 bp	47,657	18,163 bp	2 (chr1+chr2)	2,947,270 bp (chr1) 1,049,609 bp (chr2)	3,996,879 bp	245 x	47.9% (chr1) 47.0% (chr2
L6G	MB#956	SAMN14944070	CP053802 (chr1) CP053803 (chr2)	1,134,851,816 bp	70,295	16,144 bp	2 (chr1+chr2)	2,829,740 bp (chr1) 1,195,206 bp (chr2)	4,024,946 bp	225 x	48.1% (chr1) 46.3% (chr2)
SL6Y	MB#953	SAMN14944072	CP053804 (chr1) CP053805 (chr2)	826,125,289 bp	48,156	17,155 bp	2 (chr1+chr2)	2,900,064 bp (chr1) 1,143,503 bp (chr2)	4,043,567 bp	175 x	47.8% (chr1) 46.7% (chr2)
SP6G	MB#964	SAMN14944080	CP053806 (chr1) CP053807 (chr2)	874,659,808 bp	47,923	18,251 bp	2 (chr1+chr2)	2,947,818 bp (chr1) 1,229,641 bp (chr2)	4,177,459 bp	154 x	47,8% (chr1) 46.7% (chr2)
SP7G	MB#952	SAMN14944081	CP053808 (chr1) CP053809 (chr2)	1,339,636,117 bp	73,934	18,119 bp	3 (chr1+chr2)	2,868,038 bp (chr1) 1,217,051 bp (chr2)	4,085,089 bp	217 x	48.0% (chr1) 46.5% (chr2)

W6G	MB#354	SAMN14944082	CP053810 (chr1) CP053811 (chr2) CP053812 (plasmid)	1,247,754,152 bp	69,891	17,852 bp	3 (chr1+chr2+ plasmid)	2,967,360 bp (chr1) 1,202,132 bp (chr2) 306,484 bp (plas)	4,475,976 bp	192 x	47.8% (chr1) 46.6% (chr2) 42.2% (plas)
W7G	MB#962	SAMN14944083	CP053813 (chr1) CP053814 (chr2) CP053815 (plasmid)	1,490,677,399 bp	86,303	17,272 bp	3 (chr1+chr2+ plasmid)	2,967,793 bp (chr1) 1,202,155 bp (chr2) 306,481 bp (plas)	4,476,429 bp	223 x	47.8% (chr1) 46.6% (chr2) 42.2% (plas)
E7G	MB#963	SAMN14944111	CP053822 (chr1) CP053823 (chr2) CP053824 (plasmid)	1,477,482,932 bp	83,856	17,619 bp	3 (chr1+chr2+ plasmid)	3,040,343 bp (chr1) 1,199,805 bp (chr2) 80,726 bp (plas)	4,320,874 bp	194 x	47.8% (chr1) 46.5% (chr2) 46.3% (plas)
SA7G	MB#959	SAMN14944112	CP053816 (chr1) CP053817 (chr2) CP053818 (plasmid 1) CP053819 (plasmid 2)	1,689,606,494 bp	98,160	17,212 bp	4 (chr1+chr2+ plasmid 1 + 2)	3,004,584 bp (chr1) 1,182,929 bp (chr2) 79,397 bp (plas1) 80,721 bp (plas2)	4,347,631 bp	245 x	47.8% (chr1) 46.5% (chr2) 42.2% (plas1) 46.3% (plas2)
SA10G	MB#5539	SAMN14944113	CP053820 (chr1) CP053821 (chr2)	1,643,277,942 bp	91,387	17,981 bp	2 (chr1+chr2)	3,010,605 bp (chr1) 1,166,701 bp (chr2)	4,177,306 bp	269 x	47.8% (chr1) 46.6% (chr2)

[#]Data based on (69).

Supplementary table 4.3 BlastP-predicted T6SS effector and immunity proteins[1] of the environmental *V. cholerae* isolates.

Clade	Strain name	Major cluster E/I	Aux cluster 1 *vgrG1*	Aux cluster 1 E/I	Aux cluster 2 E/I	Aux cluster 3 E/I
Pandemic	**A1552**	Encoded protein with 1018 aa; evolved VgrG3 with predicted peptidoglycan hydrolase domain; A-type.	Encoded protein with 1164 aa; evolved VgrG1 with an actin crosslinking domain (ACD).	Encoded protein with 642 aa; lipase class 3; A-type.	Encoded protein with 1086 aa; VasX; pore forming activity[2]; A-type.	Encoded protein with 224 aa; papain-like NlpC/P60 peptidase[3]; A-type.
D	**SA10G***	Encoded protein with 995 aa; evolved VgrG3 with predicted peptidoglycan hydrolase and pesticin-like lysozyme domains; J-type.	Encoded protein with 1164 aa; evolved VgrG1 with an actin crosslinking domain (ACD).	Encoded protein with 710 aa; uncharacterized alpha/beta hydrolase domain (DUF2235); C-type.	Encoded protein with 1194 aa; predicted lysin motif (binds peptidoglycan), but also with pore formation activity[4]; E-type.	Aux cluster 3 is absent.
	E7G*	Encoded protein with 995 aa; evolved VgrG3 with predicted peptidoglycan hydrolase and pesticin-like lysozyme domains; J-type.	Encoded protein with 1164 aa; evolved VgrG1 with an actin crosslinking domain (ACD).	Encoded protein with 710 aa; uncharacterized alpha/beta hydrolase domain (DUF2235); C-type.	Encoded protein with 1197 aa; predicted lysin motif (binds peptidoglycan), but also with pore formation activity[4]; E-type.	Aux cluster 3 is absent.
	SA7G*	Encoded protein with 995 aa; evolved VgrG3 with predicted peptidoglycan hydrolase and pesticin-like lysozyme domains; J-type.	Encoded protein with 1164 aa; evolved VgrG1 with an actin crosslinking domain (ACD).	Encoded protein with 710 aa; uncharacterized alpha/beta hydrolase domain (DUF2235); C-type.	Encoded protein with 1197 aa; predicted lysin motif (binds peptidoglycan), but also with pore formation activity[4]; E-type.	Aux cluster 3 is absent.
C	**SL6Y**	Encoded protein with 977 aa; evolved VgrG3 with a predicted peptidoglycan hydrolase and phage-like lysozyme domains. E-type.	Encoded protein with 1164 aa; evolved VgrG1 with an actin crosslinking domain (ACD).	Encoded protein with 710 aa; uncharacterized alpha/beta hydrolase domain (DUF2235); C-type.	Encoded protein with 1086 aa; 98.3% identical to A1552's VasX; pore forming activity[2]; A-type.	Aux cluster 3 is absent.
	SP6G*	Encoded protein with 1052 aa; evolved VgrG3 with a predicted peptidoglycan hydrolase domain; H-type.	Encoded protein with 1164 aa; evolved VgrG1 with an actin crosslinking domain (ACD).	Encoded protein with 710 aa; uncharacterized alpha/beta hydrolase domain (DUF2235); C-type.	Encoded protein with 1087 aa; 94.8% identical to A1552's VasX; pore forming activity[2]; A-type.	Aux cluster 3 is absent.
	L6G*	Encoded protein with 989 aa; evolved VgrG3 with a	Encoded protein with 1167 aa;	Encoded protein with 710 aa; uncharacterized alpha/beta	Encoded protein with 1194 aa; predicted lysin motif (binds	Aux cluster 3 is absent.

		predicted peptidoglycan hydrolase and lambda phage-like lysozyme domains; K-type.	evolved VgrG1 with an actin crosslinking domain (ACD).	hydrolase domain (DUF2235); C-type.	peptidoglycan), but also with pore formation activity[4]; E-type.	
	W6G*	Encoded protein with 1052 aa; evolved VgrG3 with a predicted peptidoglycan hydrolase domain; H-type.	Encoded protein with 1164 aa; evolved VgrG1 with an actin crosslinking domain (ACD).	Encoded protein with 710 aa; uncharacterized alpha/beta hydrolase domain (DUF2235); C-type.	Encoded protein with 1088 aa; 94.1% identical to A1552's VasX; pore forming activity[2]; A-type.	Aux cluster 3 is absent.
	W7G*	Encoded protein with 1052 aa; evolved VgrG3 with a predicted peptidoglycan hydrolase domain; H-type.	Encoded protein with 1164 aa; evolved VgrG1 with an actin crosslinking domain (ACD).	Encoded protein with 710 aa; uncharacterized alpha/beta hydrolase domain (DUF2235); C-type.	Encoded protein with 1088 aa; 94.1% identical to A1552's VasX; pore forming activity[2]; A-type.	Aux cluster 3 is absent.
	SP7G**	Encoded protein with 439 aa, independent from VgrG3 (structural only); cell adhesion function[4]; B-type.	Encoded protein with 1164 aa; evolved VgrG1 with an actin crosslinking domain (ACD).	Encoded protein with 710 aa; uncharacterized alpha/beta hydrolase domain (DUF2235); C-type.	Encoded protein with 1194 aa; predicted lysin motif (binds peptidoglycan), but also with pore formation activity[4]; E-type.	Aux cluster 3 is absent.
A	**W10G**	Encoded protein with 1018 aa; evolved VgrG3 with predicted peptidoglycan hydrolase domain; A-type.	Encoded protein with 684 aa; structural VgrG1 protein.	PacBio sequencing artifact in the effector gene (HPY12_07370). Region was Sanger sequenced and manually corrected before translation to allow typing. Encoded protein with 710 aa; uncharacterized alpha/beta hydrolase domain (DUF2235); C-type.	PacBio sequencing artifact in the effector gene (HPY12_14010). Region was Sanger sequenced and manually corrected before translation to allow typing. Encoded protein with 1086 aa; VasX; pore forming activity[2]; A-type.	Aux cluster 3 is absent.
B	**SA3G**	Encoded protein with 1018 aa; evolved VgrG3 with predicted peptidoglycan hydrolase domain; A-type.	Encoded protein with 684 aa; structural VgrG1 protein.	Encoded protein with 710 aa; uncharacterized alpha/beta hydrolase domain (DUF2235); C-type.	Encoded protein with 1087 aa; 94.3% identical to A1552's VasX; pore forming activity[2]; A-type.	Aux cluster 3 is absent.
	SA5Y	Encoded protein with 995 aa; evolved VgrG3 with predicted peptidoglycan hydrolase and pesticin-like lysozyme domains; J-type.	Encoded protein with 684 aa; structural VgrG1 protein.	Encoded protein with 710 aa; uncharacterized alpha/beta hydrolase domain (DUF2235); C-type.	Encoded protein with 1088 aa; 93.7% identical to A1552's VasX; pore forming activity[2]; A-type.	Aux cluster 3 is absent.
	SL4G*	Encoded protein with 995 aa; evolved VgrG3 with predicted	Encoded protein with 684 aa;	Encoded protein with 710 aa; uncharacterized alpha/beta	Encoded protein with 1088 aa; 93.7% identical to A1552's	Aux cluster 3 is absent.

		peptidoglycan hydrolase and pesticin-like lysozyme domains; J-type.	structural VgrG1 protein.	hydrolase domain (DUF2235); C-type.	VasX; pore forming activity[2]; A-type.	
	SL5Y*	Encoded protein with 995 aa; evolved VgrG3 with predicted peptidoglycan hydrolase and pesticin-like lysozyme domains; J-type.	Encoded protein with 684 aa; structural VgrG1 protein.	Encoded protein with 710 aa; uncharacterized alpha/beta hydrolase domain (DUF2235); C-type.	Encoded protein with 1088 aa; 93.7% identical to A1552's VasX; pore forming activity[2]; A-type.	Aux cluster 3 is absent.
	SO5Y*	Encoded protein with 1022 aa; evolved VgrG3 with predicted peptidoglycan hydrolase and muraidase (N-acetylmuramidase) domains; I-type.	Encoded protein with 684 aa; structural VgrG1 protein.	Encoded protein with 710 aa; uncharacterized alpha/beta hydrolase domain (DUF2235); C-type.	Encoded protein with 1220 aa; predicted lysin motif (binds peptidoglycan); D-type.	Aux cluster 3 is absent.
Non-pandemic	**ATCC25872**	Encoded protein with 1018 aa; evolved VgrG3 with predicted peptidoglycan hydrolase domain; A-type.	Encoded protein with 1164 aa; evolved VgrG1 with an actin crosslinking domain (ACD).	Encoded protein with 642 aa; lipase class 3; A-type.	Encoded protein with 1086 aa; VasX; pore forming activity[2]; A-type.	Encoded protein with 224 aa; papain-like NlpC/P60 peptidase[3]; A-type.

[1] Typing according to (46) & (86).
[2] According to (55).
[3] According to (59).
[4] According to (46).
* In PGAP annotation, *vgrG3* gene was annotated without the effector portion. Prokka annotation used instead.
** Strain SP7G (clade C) contains an additional T6SS cluster, called Aux 6, which has the same gene order as auxiliary clusters 1 and 2 (see Figure 4.4).

Supplementary table 4.4 Matrices of % identity (percentage of residues which are identical) among VgrG1-ACD proteins in T6SS auxiliary cluster 1 from clade C and D environmental *V. cholerae* strains and pandemic strain A1552.

VgrG1-ACD full protein sequence										
	A1552	SA10G	E7G	SA7G	SL6Y	SP6G	L6G	W7G	W6G	SP7G
A1552		98.194	98.194	98.194	98.624	98.796	97.256	98.452	98.452	98.366
SA10G			100	100	98.882	98.882	97.17	99.054	99.054	98.968
E7G				100	98.882	98.882	97.17	99.054	99.054	98.968
SA7G					98.882	98.882	97.17	99.054	99.054	98.968
SL6Y						98.968	97.427	98.968	98.968	98.968
SP6G							97.427	99.14	99.14	99.054
L6G								97.684	97.684	97.427
W7G									100	99.398
W6G										99.398
SP7G										

Actin cross-linking domain (ACD) protein sequence only										
	A1552	SA10G	E7G	SA7G	SL6Y	SP6G	L6G	W7G	W6G	SP7G
A1552		99.582	99.582	99.582	99.582	99.582	99.582	99.582	99.582	99.374
SA10G			100	100	100	99.582	100	100	100	99.791
E7G				100	100	99.582	100	100	100	99.791
SA7G					100	99.582	100	100	100	99.791
SL6Y						99.582	100	100	100	99.791
SP6G							99.582	99.582	99.582	99.374
L6G								100	100	99.791
W7G									100	99.791
W6G										99.791
SP7G										

Green: 100% identity
Yellow: 90% - 99.99% identity

Supplementary table 4.5 Matrices of % identity (percentage of residues that are identical) among the effector and immunity proteins harbored in the T6SS large cluster from the environmental *V. cholerae* strains and the pandemic strain A1552.

Large cluster – Effector																
	A1552	SA10G	E7G	SA7G	SL6Y	SP6G	L6G	W7G	W6G	SP7G	W10G	SA3G	SA5Y	SL4G	SL5Y	SO5Y
A1552		8.9	8.9	8.9	7.5	6.7	5.7	6.7	6.7	5.3	100	99.5	8.9	8.8	8.9	9.8
SA10G			100	100	13.9	8.4	8.2	8.4	8.4	7.4	8.9	8.9	98.9	98.9	98.9	10.4
E7G				100	13.9	8.4	8.2	8.4	8.4	7.4	8.9	8.9	98.9	98.9	98.9	10.4
SA7G					13.9	8.4	8.2	8.4	8.4	7.4	8.9	8.9	98.9	98.9	98.9	10.4
SL6Y						8.6	10.6	8.6	8.6	5.6	7.5	7.5	13.8	13.8	13.8	6.5
SP6G							9.2	100	100	5.7	6.7	6.7	8.3	8.3	8.4	10.8
L6G								9.2	9.2	5.4	5.7	5.7	8.7	8.6	8.4	8.5
W7G									100	5.7	6.7	6.7	8.3	8.3	8.4	10.8
W6G										5.7	6.7	6.7	8.3	8.3	8.4	10.8
SP7G											5.3	5.3	7.6	7.6	7.4	5.3
W10G												99.5	8.9	8.8	8.9	9.8
SA3G													8.9	8.8	8.9	9.8
SA5Y														100	100	10.8
SL4G															100	10.8
SL5Y																10.9
SO5Y																

Large cluster – Immunity																
	A1552	SA10G	E7G	SA7G	SL6Y	SP6G	L6G	W7G	W6G	SP7G	W10G	SA3G	SA5Y	SL4G	SL5Y	SO5Y
A1552		11.7	11.7	11.7	7.7	8.1	12.0	8.1	8.1	6.4	100	100	11.7	11.7	11.7	15.8
SA10G			100	100	10.8	6.8	8.4	6.8	6.8	10.9	11.7	11.7	100	100	100	10.5
E7G				100	10.8	6.8	8.4	6.8	6.8	10.9	11.7	11.7	100	100	100	10.5
SA7G					10.8	6.8	8.4	6.8	6.8	10.9	11.7	11.7	100	100	100	10.5
SL6Y						8.5	15.8	8.5	8.5	8.8	7.7	7.7	10.8	10.8	10.8	10.1
SP6G							7.0	100	100	8.4	8.1	8.1	6.8	6.8	6.8	9.7
L6G								7.0	7.0	8.5	12.0	12.0	8.4	8.4	8.4	11.2
W7G									100	8.4	8.1	8.1	6.8	6.8	6.8	9.7
W6G										8.4	8.1	8.1	6.8	6.8	6.8	9.7
SP7G											6.4	6.4	10.9	10.9	10.9	7.5
W10G												100	11.7	11.7	11.7	15.8
SA3G													11.7	11.7	11.7	15.8
SA5Y														100	100	10.5
SL4G															100	10.5
SL5Y																10.5
SO5Y																

Green: 100% identity (same family and subfamily)
Yellow: 90% - 99.99% identity (same family, with some polymorphism)
Orange: 30% - 90% identity (same family, with considerable polymorphism)
Red: <30% identity (different families)

Supplementary table 4.6 Matrices of % identity (percentage of residues that are identical) among the effector and immunity proteins harbored in the T6SS auxiliary cluster 1 from the environmental *V. cholerae* strains and the pandemic strain A1552.

Auxiliary cluster 1 – Effector																
	A1552	SA10G	E7G	SA7G	SL6Y	SP6G	L6G	W7G	W6G	SP7G	W10G	SA3G	SA5Y	SL4G	SL5Y	SO5Y
A1552		15.0	15.0	15.0	15.5	15.0	15.5	15.8	15.8	15.8	14.6	15.9	15.9	15.9	14.9	15.3
SA10G			100	100	90.4	99.7	90.6	94.6	94.6	94.6	91.3	94.5	94.5	94.4	92.5	98.9
E7G				100	90.4	99.7	90.6	94.6	94.6	94.6	91.3	94.5	94.5	94.4	92.5	98.9
SA7G					90.4	99.7	90.6	94.6	94.6	94.6	91.3	94.5	94.5	94.4	92.5	98.9
SL6Y						90.4	99.9	92.9	92.9	92.9	87.0	92.8	92.8	92.7	94.5	91.1
SP6G							90.6	94.4	94.4	94.3	91.3	94.2	94.2	94.1	92.3	98.9
L6G								93.1	93.1	93.1	87.2	92.9	92.9	92.8	94.6	91.3
W7G									100	100	90.3	99.6	99.6	99.4	91.0	94.6
W6G										100	90.3	99.6	99.6	99.4	91.0	94.6
SP7G											90.3	99.6	99.6	99.4	91.0	94.6
W10G												90.1	90.1	89.9	89.9	91.9
SA3G													100	99.9	90.8	94.5
SA5Y														99.9	90.83	94.5
SL4G															90.7	94.4
SL5Y																92.7
SO5Y																

Auxiliary cluster 1 – Immunity																
	A1552	SA10G	E7G	SA7G	SL6Y	SP6G	L6G	W7G	W6G	SP7G	W10G	SA3G	SA5Y	SL4G	SL5Y	SO5Y
A1552		15.5	15.5	15.5	15.2	15.5	15.2	15.9	15.9	15.9	14.0	15.9	15.9	15.9	15.1	15.1
SA10G			100	100	72.1	99.6	71.7	85.7	85.7	85.7	78.3	85.7	85.7	85.7	76.2	97.5
E7G				100	72.1	99.6	71.7	85.7	85.7	85.7	78.3	85.7	85.7	85.7	76.2	97.5
SA7G					72.1	99.6	71.7	85.7	85.7	85.7	78.3	85.7	85.7	85.7	76.2	97.5
SL6Y						72.1	99.6	78.7	78.7	78.7	65.6	78.7	78.7	78.7	91.4	72.5
SP6G							71.7	85.7	85.7	85.7	77.9	85.7	85.7	85.7	75.8	97.1
L6G								78.3	78.3	78.3	65.2	78.3	78.3	78.3	90.9	72.1
W7G									100	100	71.3	100	100	100	73.0	85.2
W6G										100	71.3	100	100	100	73.0	85.2
SP7G											71.3	100	100	100	73.0	85.2
W10G												71.3	71.3	71.3	70.2	79.2
SA3G													100	100	73.0	85.2
SA5Y														100	73.0	85.2
SL4G															73.0	85.2
SL5Y																76.3
SO5Y																

Green: 100% identity (same family and subfamily)
Yellow: 90% - 99.99% identity (same family, with some polymorphism)
Orange: 30% - 90% identity (same family, with considerable polymorphism)
Red: <30% identity (different families)

Supplementary table 4.7 Matrices of % identity (percentage of residues that are identical) among the effector and immunity proteins harbored in the T6SS auxiliary cluster 2 from the environmental *V. cholerae* strains and the pandemic strain A1552.

Auxiliary cluster 2 – Effector																
	A1552	SA10G	E7G	SA7G	SL6Y	SP6G	L6G	W7G	W6G	SP7G	W10G	SA3G	SA5Y	SL4G	SL5Y	SO5Y
A1552		14.6	14.4	14.4	98.3	94.8	14.8	94.0	94.0	14.8	96.5	94.3	93.6	93.6	93.6	12.7
SA10G			95.2	95.2	14.5	15.3	97.6	15.2	15.2	97.5	14.9	15.0	15.2	15.2	15.2	20.2
E7G				100	14.3	15.0	95.1	15.0	15.0	95.0	14.6	14.7	15.1	15.1	15.1	19.8
SA7G					14.3	15.0	95.1	15.0	15.0	95.0	14.6	14.7	15.1	15.1	15.1	19.8
SL6Y						94.1	14.7	93.7	93.5	14.7	95.7	93.8	93.7	93.7	93.7	12.8
SP6G							15.4	96.1	96.1	15.4	94.7	98.8	95.0	95.0	95.0	12.7
L6G								15.3	15.3	99.9	15.0	15.2	15.4	15.4	15.4	19.8
W7G									100	15.3	93.8	94.8	99.0	99.0	99.0	13.0
W6G										15.3	93.8	94.8	99.0	99.0	99.0	13.0
SP7G											15.0	15.2	15.4	15.4	15.4	19.8
W10G												94.0	93.4	93.4	93.4	12.8
SA3G													94.0	94.0	94.0	12.4
SA5Y														100	100	13.0
SL4G															100	13.0
SL5Y																13.0
SO5Y																

Auxiliary cluster 2 – Immunity																
	A1552	SA10G	E7G	SA7G	SL6Y	SP6G	L6G	W7G	W6G	SP7G	W10G	SA3G	SA5Y	SL4G	SL5Y	SO5Y
A1552		13.2	12.9	12.9	99.6	72.3	12.9	68.7	68.7	12.9	83.9	71.9	68.3	68.3	68.2	12.4
SA10G			73.1	73.1	13.2	13.6	92.5	12.9	12.9	92.2	12.5	13.6	12.9	12.9	12.9	9.4
E7G				100	12.9	11.2	73.9	11.9	11.9	73.9	12.2	11.5	11.9	11.9	11.9	8.1
SA7G					12.9	11.2	73.9	11.9	11.9	73.9	12.2	11.5	11.9	11.9	11.9	8.1
SL6Y						71.9	12.9	68.3	68.3	12.9	83.5	71.5	67.9	67.9	67.9	12.4
SP6G							12.5	80.1	80.1	12.5	74.0	99.6	80.5	80.5	80.5	12.1
L6G								12.5	12.5	99.7	12.2	12.5	12.5	12.5	12.5	9.9
W7G									100	12.5	67.5	79.7	99.2	99.2	99.2	11.0
W6G										12.5	67.5	79.7	99.2	99.2	99.2	11.0
SP7G											12.2	12.5	12.5	12.5	12.5	9.9
W10G												73.6	67.5	67.5	67.5	11.2
SA3G													80.1	80.1	80.1	12.1
SA5Y														100	100	11.0
SL4G															100	11.0
SL5Y																11.0
SO5Y																

Green: 100% identity (same family and subfamily)
Yellow: 90% - 99.99% identity (same family, with some polymorphism)
Orange: 30% - 90% identity (same family, with considerable polymorphism)
Red: <30% identity (different families)

4.11 REFERENCES

1. Ali,M., Nelson,A.R., Lopez,A.L. and Sack,D.A. (2015) Updated global burden of cholera in endemic countries. *PLoS Negl. Trop. Dis.*, **9**, e0003832.

2. Faruque,S.M., Albert,M.J. and Mekalanos,J.J. (1998) Epidemiology, genetics, and ecology of toxigenic *Vibrio cholerae*. *Microbiol. Mol. Biol. Rev.*, **62**, 1301–14.

3. World Health Organization (WHO) (2006) Communicable diseases following natural disasters Risk assessment and priority interventions.

4. Clemens,J.D., Nair,G.B., Ahmed,T., Qadri,F. and Holmgren,J. (2017) Cholera. *Lancet*, **390**, 1539–1549.

5. Cottingham,K.L., Chiavelli,D.A. and Taylor,R.K. (2003) Environmental microbe and human pathogen: the ecology and microbiology of *Vibrio cholerae*. *Front Ecol Environm*, **1**, 80–86.

6. Johnson,J.A., Salles,C.A., Panigrahi,P., Albert,M.J., Wright,A.C., Johnson,R.J. and Morris,J.G. (1994) *Vibrio cholerae* O139 synonym Bengal is closely related to *Vibrio cholerae* El Tor but has important differences. *Infect. Immun.*, **62**, 2108–2110.

7. Kirn,T.J., Jude,B.A. and Taylor,R.K. (2005) A colonization factor links *Vibrio cholerae* environmental survival and human infection. *Nature*, **438**, 863–866.

8. Dziejman,M., Balon,E., Boyd,D., Fraser,C.M., Heidelberg,J.F. and Mekalanos,J.J. (2002) Comparative genomic analysis of *Vibrio cholerae*: Genes that correlate with cholera endemic and pandemic disease. *Proc. Natl. Acad. Sci.*, **99**, 1556–1561.

9. Dziejman,M., Serruto,D., Tam,V.C., Sturtevant,D., Diraphat,P., Faruque,S.M., Rahman,M.H., Heidelberg,J.F., Decker,J., Li,L., *et al.* (2005) Genomic characterization of non-O1, non-O139 *Vibrio cholerae* reveals genes for a type III secretion system. *Proc. Natl. Acad. Sci.*, **102**, 3465–3470.

10. Singh,D. V, Matte,M.H., Matte,G.R., Jiang,S., Sabeena,F., Shukla,B.N., Sanyal,S.C., Huq,A. and Colwell,R.R. (2001) Molecular analysis of *Vibrio cholerae* O1, O139, non-O1, and non-O139 strains: clonal relationships between clinical and environmental Isolates. *Appl. Environ. Microbiol.*, **67**, 910–921.

11. Onifade,T.J.M., Hutchinson,R., van Zile,K., Bodager,D., Baker,R. and Blackmore,C. (2011) Toxin producing *Vibrio cholerae* O75 outbreak, United States, March to April 2011. *Eurosurveillance*, **16**, pii=19870.

12. Deshayes,S., Daurel,C., Cattoir,V., Parienti,J.J., Quilici,M.L. and de La Blanchardière,A. (2015) Non-O1, non-O139 *Vibrio cholerae* bacteraemia: case report and literature review. *Springerplus*, **4**.

13. Islam,M.S., Alam,M.J. and Neogi,P.K.B. (1992) Seasonality and toxigenicity of *Vibrio cholerae* non-01 isolated from different components of pond ecosystems of Dhaka City, Bangladesh. *World J. Microbiol. Biotechnol.*, **8**, 160–163.

14. Hasan,N.A., Rezayat,T., Blatz,P.J., Choi,S.Y., Griffitt,K.J., Rashed,S.M., Huq,A., Conger,N.G., Colwell,R.R. and Jay Grimes,D. (2015) Nontoxigenic *Vibrio cholerae* Non-O1/O139 isolate from a case of human gastroenteritis in the U.S. Gulf Coast. *J. Clin. Microbiol.*, **53**, 9–14.

15. Harris,J.B., LaRocque,R.C., Qadri,F., Ryan,E.T. and Calderwood,S.B. (2012) Cholera. *Lancet*, **379**, 2466–2476.

16. Chun,J., Grim,C.J., Hasan,N.A., Lee,J.H., Choi,S.Y., Haley,B.J., Taviani,E., Jeon,Y.-S., Kim,D.W., Lee,J.-H., *et al.* (2009) Comparative genomics reveals mechanism for short-term and long-term clonal transitions in pandemic *Vibrio cholerae*. *Proc. Natl. Acad. Sci.*, **106**, 15442–15447.

17. Mutreja,A., Kim,D.W., Thomson,N.R., Connor,T.R., Lee,J.H., Kariuki,S., Croucher,N.J., Choi,S.Y., Harris,S.R., Lebens,M., *et al.* (2011) Evidence for several waves of global transmission in the seventh cholera pandemic. *Nature*, **477**, 462–465.

18. Weill,F.-X., Domman,D., Njamkepo,E., Tarr,C., Rauzier,J., Fawal,N., Keddy,K.H., Salje,H., Moore,S.,

Mukhopadhyay,A.K., *et al.* (2017) Genomic history of the seventh pandemic of cholera in Africa. *Science*, **358**, 785–789.

19. Domman,D., Quilici,M.-L., Dorman,M.J., Njamkepo,E., Mutreja,A., Mather,A.E., Delgado,G., Morales-Espinosa,R., Grimont,P.A.D., Lizarraga-Partida,M.L., *et al.* (2017) Integrated view of *Vibrio cholerae* in the Americas. *Science*, **358**, 789–793.

20. Weill,F.X., Domman,D., Njamkepo,E., Almesbahi,A.A., Naji,M., Nasher,S.S., Rakesh,A., Assiri,A.M., Sharma,N.C., Kariuki,S., *et al.* (2019) Genomic insights into the 2016–2017 cholera epidemic in Yemen. *Nature*, **565**, 230–233.

21. Waldor,M.K. and Mekalanos,J.J. (1996) Lysogenic conversion by a filamentous phage encoding cholera toxin. *Science*, **272**, 1910–1914.

22. Taylor,R.K., Miller,V.L., Furlong,D.B. and Mekalanos,J.J. (1987) Use of *phoA* gene fusions to identify a pilus colonization factor coordinately regulated with cholera toxin. *Proc. Natl. Acad. Sci.*, **84**, 2833–2837.

23. Faruque,S.M., Chowdhury,N., Kamruzzaman,M., Dziejman,M., Rahman,M.H., Sack,D.A., Nair,G.B. and Mekalanos,J.J. (2004) Genetic diversity and virulence potential of environmental *Vibrio cholerae* population in a cholera-endemic area. *Proc. Natl. Acad. Sci.*, **101**, 2123–2128.

24. Shapiro,B.J., Levade,I., Kovacikova,G., Taylor,R.K. and Almagro-Moreno,S. (2016) Origins of pandemic *Vibrio cholerae* from environmental gene pools. *Nat. Microbiol.*, **2**, 16240.

25. Rivera,I.N.G., Chun,J., Huq,A., Sack,R.B. and Colwell,R.R. (2001) Genotypes associated with virulence in environmental isolates of *Vibrio cholerae*. *Appl. Environ. Microbiol.*, **67**, 2421–2429.

26. Gennari,M., Ghidini,V., Caburlotto,G. and Lleo,M.M. (2012) Virulence genes and pathogenicity islands in environmental *Vibrio* strains nonpathogenic to humans. *FEMS Microbiol. Ecol.*, **82**, 563–573.

27. Faruque,S.M., Kamruzzaman,M., Meraj,I.M., Chowdhury,N., Nair,G.B., Sack,R.B., Colwell,R.R. and Sack,D.A. (2003) Pathogenic potential of environmental *Vibrio cholerae* strains carrying genetic variants of the toxin-coregulated pilus pathogenicity island. *Infect. Immun.*, **71**, 1020–1025.

28. Bernardy,E.E., Turnsek,M.A., Wilson,S.K., Tarr,C.L. and Hammer,B.K. (2016) Diversity of clinical and environmental isolates of *Vibrio cholerae* in natural tansformation and contact-dependent bacterial killing indicative of type VI secretion system activity. *Appl. Environ. Microbiol.*, **82**, 2833–2842.

29. Aldova,E., Laznickova,K., Stepankova,E. and Lietava,J. (1968) Isolation of nonagglutinable vibrios from an enteritis outbreak in Czechoslovakia. *J. Infect. Dis.*, **118**, 25–31.

30. Boyd,E.F. and Waldor,M.K. (2002) Evolutionary and functional analyses of variants of the toxin-coregulated pilus protein TcpA from toxigenic *Vibrio cholerae* non-O1/nonO139 serogroup isolates. *Microbiology*, **148**, 1655–1666.

31. Ho,B.T., Dong,T.G. and Mekalanos,J.J. (2014) A view to a kill: the bacterial type VI secretion system. *Cell Host Microbe*, **15**, 9–21.

32. Cianfanelli,F.R., Monlezun,L. and Coulthurst,S.J. (2016) Aim, load, fire: the type VI secretion system, a bacterial nanoweapon. *Trends Microbiol.*, **24**, 51–62.

33. Galán,J.E. and Waksman,G. (2018) Protein-injection machines in bacteria. *Cell*, **172**, 1306–1318.

34. Taylor,N.M.I., van Raaij,M.J. and Leiman,P.G. (2018) Contractile injection systems of bacteriophages and related systems. *Mol. Microbiol.*, **108**, 6–15.

35. Cherrak,Y., Flaugnatti,N., Durand,E., Journet,L. and Cascales,E. (2019) Structure and regulation of the type VI secretion system. *Microbiol. Spectr.*, **7**, PSIB-0031-2019.

36. Zoued,A., Brunet,Y.R., Durand,E., Aschtgen,M.S., Logger,L., Douzi,B., Journet,L., Cambillau,C. and Cascales,E. (2014) Architecture and assembly of the type VI secretion system. *Biochim. Biophys. Acta - Mol. Cell Res.*, **1843**, 1664–1673.

37. Flaugnatti,N., Rapisarda,C., Rey,M., Beauvois,S.G., Nguyen,V.A., Canaan,S., Durand,E., Chamot-Rooke,J., Cascales,E., Fronzes,R., *et al.* (2020) Structural basis for loading and inhibition of a bacterial T6SS phospholipase effector by the VgrG spike. *EMBO J.*, **39**, e104129.

38. Pukatzki,S., Ma,A.T., Sturtevant,D., Krastins,B., Sarracino,D., Nelson,W.C., Heidelberg,J.F. and Mekalanos,J.J. (2006) Identification of a conserved bacterial protein secretion system in *Vibrio cholerae* using the *Dictyostelium* host model system. *Proc. Natl. Acad. Sci.*, **103**, 1528–33.

39. Basler,M., Pilhofer,M., Henderson,G.P., Jensen,G.J. and Mekalanos,J.J. (2012) Type VI secretion requires a dynamic contractile phage tail-like structure. *Nature*, **483**, 182–186.

40. Bröms,J.E., Ishikawa,T., Wai,S.N. and Sjöstedt,A. (2013) A functional VipA-VipB interaction is required for the type VI secretion system activity of *Vibrio cholerae* O1 strain A1552. *BMC Microbiol.*, **13**.

41. Kube,S., Kapitein,N., Zimniak,T., Herzog,F., Mogk,A. and Wendler,P. (2014) Structure of the VipA/B type VI secretion complex suggests a contraction-state-specific recycling mechanism. *Cell Rep.*, **8**, 20–30.

42. Hood,R.D., Singh,P., Hsu,F.S., Güvener,T., Carl,M.A., Trinidad,R.R.S., Silverman,J.M., Ohlson,B.B., Hicks,K.G., Plemel,R.L., *et al.* (2010) A type VI secretion system of *Pseudomonas aeruginosa* targets a toxin to bacteria. *Cell Host Microbe*, **7**, 25–37.

43. Russell,A.B., Peterson,S.B. and Mougous,J.D. (2014) Type VI secretion system effectors: poisons with a purpose. *Nat. Rev. Microbiol.*, **12**, 137–148.

44. Pukatzki,S., McAuley,S.B. and Miyata,S.T. (2009) The type VI secretion system: translocation of effectors and effector-domains. *Curr. Opin. Microbiol.*, **12**, 11–17.

45. Hachani,A., Wood,T.E. and Filloux,A. (2016) Type VI secretion and anti-host effectors. *Curr. Opin. Microbiol.*, **29**, 81–93.

46. Unterweger,D., Miyata,S.T., Bachmann,V., Brooks,T.M., Mullins,T., Kostiuk,B., Provenzano,D. and Pukatzki,S. (2014) The *Vibrio cholerae* type VI secretion system employs diverse effector modules for intraspecific competition. *Nat. Commun.*, **5**.

47. Dong,T.G., Ho,B.T., Yoder-Himes,D.R. and Mekalanos,J.J. (2013) Identification of T6SS-dependent effector and immunity proteins by Tn-seq in *Vibrio cholerae*. *Proc. Natl. Acad. Sci.*, **110**, 2623–2628.

48. Brooks,T.M., Unterweger,D., Bachmann,V., Kostiuk,B. and Pukatzki,S. (2013) Lytic activity of the *Vibrio cholerae* type VI secretion toxin VgrG-3 is inhibited by the antitoxin TsaB. *J. Biol. Chem.*, **288**, 7618–7625.

49. Russell,A.B., Hood,R.D., Bui,N.K., Leroux,M., Vollmer,W. and Mougous,J.D. (2011) Type VI secretion delivers bacteriolytic effectors to target cells. *Nature*, **475**, 343–349.

50. Crisan,C. V and Hammer,B.K. (2020) The *Vibrio cholerae* type VI secretion system: toxins, regulators and consequences. *Environ. Microbiol.*, **22**, 4112–4122.

51. Zheng,J., Ho,B. and Mekalanos,J.J. (2011) Genetic analysis of anti-amoebae and anti-bacterial activities of the type VI secretion system in *Vibrio cholerae*. *PLoS One*, **6**, e23876.

52. Liang,X., Moore,R., Wilton,M., Wong,M.J.Q., Lam,L. and Dong,T.G. (2015) Identification of divergent type VI secretion effectors using a conserved chaperone domain. *Proc. Natl. Acad. Sci.*, **112**, 9106–9111.

53. Unterweger,D., Kostiuk,B., Otjengerdes,R., Wilton,A., Diaz-Satizabal,L. and Pukatzki,S. (2015) Chimeric adaptor proteins translocate diverse type VI secretion system effectors in *Vibrio cholerae*. *EMBO J.*, **34**, 2198–2210.

54. Russell,A.B., Leroux,M., Hathazi,K., Agnello,D.M., Ishikawa,T., Wiggins,P.A., Wai,S.N. and Mougous,J.D. (2013) Diverse type VI secretion phospholipases are functionally plastic antibacterial effectors. *Nature*, **496**, 508–512.

55. Miyata,S.T., Kitaoka,M., Brooks,T.M., McAuley,S.B. and Pukatzki,S. (2011) *Vibrio cholerae* requires the type VI secretion system virulence factor VasX to kill *Dictyostelium discoideum*. *Infect. Immun.*, **79**, 2941–

2949.

56. Miyata,S.T., Unterweger,D., Rudko,S.P. and Pukatzki,S. (2013) Dual expression profile of type VI secretion system immunity genes protects pandemic *Vibrio cholerae*. *PLoS Pathog.*, **9**, e1003752.

57. Shneider,M.M., Buth,S.A., Ho,B.T., Basler,M., Mekalanos,J.J. and Leiman,P.G. (2013) PAAR-repeat proteins sharpen and diversify the type VI secretion system spike. *Nature*, **500**, 350–353.

58. Altindis,E., Dong,T., Catalano,C. and Mekalanos,J.J. (2015) Secretome analysis of *Vibrio cholerae* type VI secretion system reveals a new effector-immunity pair. *MBio*, **6**, e00075-15.

59. Hersch,S.J., Watanabe,N., Stietz,M.S., Manera,K., Kamal,F., Burkinshaw,B., Lam,L., Pun,A., Li,M., Savchenko,A., *et al.* (2020) Envelope stress responses defend against type six secretion system attacks independently of immunity proteins. *Nat. Microbiol.*, **5**, 706–714.

60. Chou,S., Bui,N.K., Russell,A.B., Lexa,K.W., Gardiner,T.E., LeRoux,M., Vollmer,W. and Mougous,J.D. (2012) Structure of a peptidoglycan amidase effector targeted to Gram-negative bacteria by the type VI secretion system. *Cell Rep.*, **1**, 656–664.

61. Zhang,X.H. and Austin,B. (2005) Haemolysins in Vibrio species. *J. Appl. Microbiol.*, **98**, 1011–1019.

62. Ichinose,Y., Yamamoto,K., Nakasone,N., Tanabe,M.J., Takeda,T., Miwatani,T. and Iwanaga,M. (1987) Enterotoxicity of El Tor-like hemolysin of non-O1 *Vibrio cholerae*. *Infect. Immun.*, **55**, 1090–1093.

63. Olivier,V., Haines,G.K., Tan,Y. and Fullner Satchell,K.J. (2007) Hemolysin and the multifunctional autoprocessing RTX toxin are virulence factors during intestinal infection of mice with *Vibrio cholerae* El Tor O1 strains. *Infect. Immun.*, **75**, 5035–5042.

64. Van der Henst,C., Vanhove,A.S., Drebes Dörr,N.C., Stutzmann,S., Stoudmann,C., Clerc,S., Scrignari,T., Maclachlan,C., Knott,G. and Blokesch,M. (2018) Molecular insights into *Vibrio cholerae*'s intra-amoebal host-pathogen interactions. *Nat. Commun.*, **9**.

65. Keymer,D.P., Miller,M.C., Schoolnik,G.K. and Boehm,A.B. (2007) Genomic and phenotypic diversity of coastal *Vibrio cholerae* strains is linked to environmental factors. *Appl. Environ. Microbiol.*, **73**, 3705–3714.

66. Miller,M.C., Keymer,D.P., Avelar,A., Boehm,A.B. and Schoolnik,G.K. (2007) Detection and transformation of genome segments that differ within a coastal population of *Vibrio cholerae* strains. *Appl. Environ. Microbiol.*, **73**, 3695–3704.

67. Heidelberg,J.F., Elsen,J.A., Nelson,W.C., Clayton,R.A., Gwinn,M.L., Dodson,R.J., Haft,D.H., Hickey,E.K., Peterson,J.D., Umayam,L., *et al.* (2000) DNA sequence of both chromosomes of the cholera pathogen *Vibrio cholerae*. *Nature*, **406**, 477–483.

68. Yildiz,F.H. and Schoolnik,G.K. (1998) Role of *rpoS* in stress survival and virulence of *Vibrio cholerae*. *J. Bacteriol.*, **180**, 773–784.

69. Matthey,N., Drebes Dörr,N.C. and Blokesch,M. (2018) Long-read-based genome sequences of pandemic and environmental *Vibrio cholerae* strains. *Microbiol. Resour. Announc.*, **7**.

70. Okada,K., Iida,T., Kita-Tsukamoto,K. and Honda,T. (2005) Vibrios commonly possess two chromosomes. *J. Bacteriol.*, **187**, 752–757.

71. Unterweger,D., Kitaoka,M., Miyata,S.T., Bachmann,V., Brooks,T.M., Moloney,J., Sosa,O., Silva,D., Duran-Gonzalez,J., Provenzano,D., *et al.* (2012) Constitutive type VI secretion system expression gives *Vibrio cholerae* intra- and interspecific competitive advantages. *PLoS One*, **7**, e48320.

72. Borgeaud,S., Metzger,L.C., Scrignari,T. and Blokesch,M. (2015) The type VI secretion system of *Vibrio cholerae* fosters horizontal gene transfer. *Science*, **347**, 63–67.

73. Watve,S.S., Thomas,J. and Hammer,B.K. (2015) CytR is a global positive regulator of competence, type VI secretion, and chitinases in *Vibrio cholerae*. *PLoS One*, **10**, e01388.

74. Metzger,L.C. and Blokesch,M. (2016) Regulation of competence-mediated horizontal gene transfer in the natural habitat of *Vibrio cholerae*. *Curr. Opin. Microbiol.*, **30**, 1–7.

75. Veening,J.W. and Blokesch,M. (2017) Interbacterial predation as a strategy for DNA acquisition in naturally competent bacteria. *Nat. Rev. Microbiol.*, **15**, 621–629.

76. Metzger,L.C., Stutzmann,S., Scrignari,T., Van der Henst,C., Matthey,N. and Blokesch,M. (2016) Independent regulation of type VI secretion in *Vibrio cholerae* by TfoX and TfoY. *Cell Rep.*, **15**, 951–958.

77. Granato,E.T., Meiller-Legrand,T.A. and Foster,K.R. (2019) The evolution and ecology of bacterial warfare. *Curr. Biol.*, **29**, R521–R537.

78. Hibbing,M.E., Fuqua,C., Parsek,M.R. and Peterson,S.B. (2010) Bacterial competition: surviving and thriving in the microbial jungle. *Nat. Rev. Microbiol.*, **8**, 15–25.

79. Stubbendieck,R.M. and Straight,P.D. (2016) Multifaceted interfaces of bacterial competition. *J. Bacteriol.*, **198**, 2145–2155.

80. Pukatzki,S., Ma,A.T., Revel,A.T., Sturtevant,D. and Mekalanos,J.J. (2007) Type VI secretion system translocates a phage tail spike-like protein into target cells where it cross-links actin. *Proc. Natl. Acad. Sci.*, **104**, 15508–15513.

81. Ma,A.T., McAuley,S., Pukatzki,S. and Mekalanos,J.J. (2009) Translocation of a *Vibrio cholerae* type VI secretion effector requires bacterial endocytosis by host cells. *Cell Host Microbe*, **5**, 234–243.

82. Ma,A.T. and Mekalanos,J.J. (2010) In vivo actin cross-linking induced by *Vibrio cholerae* type VI secretion system is associated with intestinal inflammation. *Proc. Natl. Acad. Sci.*, **107**, 4365–4370.

83. Logan,S.L., Thomas,J., Yan,J., Baker,R.P., Shields,D.S., Xavier,J.B., Hammer,B.K. and Parthasarathy,R. (2018) The *Vibrio cholerae* type VI secretion system can modulate host intestinal mechanics to displace gut bacterial symbionts. *Proc. Natl. Acad. Sci. U. S. A.*, **115**, E3779–E3787.

84. Lo Scrudato,M. and Blokesch,M. (2013) A transcriptional regulator linking quorum sensing and chitin induction to render *Vibrio cholerae* naturally transformable. *Nucleic Acids Res.*, **41**, 3644–3658.

85. Jaskólska,M., Stutzmann,S., Stoudmann,C. and Blokesch,M. (2018) QstR-dependent regulation of natural competence and type VI secretion in *Vibrio cholerae*. *Nucleic Acids Res.*, **46**, 10619–10634.

86. Kirchberger,P.C., Unterweger,D., Provenzano,D., Pukatzki,S. and Boucher,Y. (2017) Sequential displacement of type VI secretion system effector genes leads to evolution of diverse immunity gene arrays in *Vibrio cholerae*. *Sci. Rep.*, **7**, 45133.

87. Matthey,N., Stutzmann,S., Stoudmann,C., Guex,N., Iseli,C. and Blokesch,M. (2019) Neighbor predation linked to natural competence fosters the transfer of large genomic regions in *Vibrio cholerae*. *Elife*, **8**, e48212.

88. Santoriello,F.J., Michel,L., Unterweger,D. and Pukatzki,S. (2019) *Vibrio cholerae* type VI secretion system auxiliary cluster 3 is a pandemic-associated mobile genetic element. *bioRxiv*, 10.1101/868539.

89. Crisan,C. V., Chande,A.T., Williams,K., Raghuram,V., Rishishwar,L., Steinbach,G., Watve,S.S., Yunker,P., Jordan,I.K. and Hammer,B.K. (2019) Analysis of *Vibrio cholerae* genomes identifies new type VI secretion system gene clusters. *Genome Biol.*, **20**, 163.

90. Labbate,M., Orata,F.D., Petty,N.K., Jayatilleke,N.D., King,W.L., Kirchberger,P.C., Allen,C., Mann,G., Mutreja,A., Thomson,N.R., *et al.* (2016) A genomic island in *Vibrio cholerae* with VPI-1 site-specific recombination characteristics contains CRISPR-Cas and type VI secretion modules. *Sci. Rep.*, **6**, 36891.

91. Wattam,A.R., Davis,J.J., Assaf,R., Boisvert,S., Brettin,T., Bun,C., Conrad,N., Dietrich,E.M., Disz,T., Gabbard,J.L., *et al.* (2017) Improvements to PATRIC, the all-bacterial bioinformatics database and analysis resource center. *Nucleic Acids Res.*, **45**, D535–D542.

92. Bishop-Lilly,K., Johnson,S., Verratti,K., Truong,L., Amy,K., Joy,A., Mokashi,V., Patrick,C. and Sozhamannan,S. (2014) Genome sequencing of 15 clinical *Vibrio* isolates, including 13 non-O1/non-

O139 serogroup strains. *Genome Announc.*, **2**, e00893-14.

93. Flaugnatti,N., Le,T.T.H., Canaan,S., Aschtgen,M.S., Nguyen,V.S., Blangy,S., Kellenberger,C., Roussel,A., Cambillau,C., Cascales,E., *et al.* (2016) A phospholipase A1 antibacterial type VI secretion effector interacts directly with the C-terminal domain of the VgrG spike protein for delivery. *Mol. Microbiol.*, **99**, 1099–1118.

94. Buist,G., Steen,A., Kok,J. and Kuipers,O.P. (2008) LysM, a widely distributed protein motif for binding to (peptido)glycans. *Mol. Microbiol.*, **68**, 838–847.

95. Durand,E., Cambillau,C., Cascales,E. and Journet,L. (2014) VgrG, Tae, Tle, and beyond: the versatile arsenal of type VI secretion effectors. *Trends Microbiol.*, **22**, 498–507.

96. McNally,L., Bernardy,E., Thomas,J., Kalziqi,A., Pentz,J., Brown,S.P., Hammer,B.K., Yunker,P.J. and Ratcliff,W.C. (2017) Killing by type VI secretion drives genetic phase separation and correlates with increased cooperation. *Nat. Commun.*, **8**, 14371.

97. Speare,L., Cecere,A.G., Guckes,K.R., Smith,S., Wollenberg,M.S., Mandel,M.J., Miyashiro,T. and Septer,A.N. (2018) Bacterial symbionts use a type VI secretion system to eliminate competitors in their natural host. *Proc. Natl. Acad. Sci. U. S. A.*, **115**, E8528–E8537.

98. Salomon,D., Klimko,J.A., Trudgian,D.C., Kinch,L.N., Grishin,N. V., Mirzaei,H. and Orth,K. (2015) Type VI secretion system toxins horizontally shared between marine bacteria. *PLoS Pathog.*, **11**, e1005128.

99. Meibom,K.L., Blokesch,M., Dolganov,N.A., Wu,C.Y. and Schoolnik,G.K. (2005) Chitin induces natural competence in *Vibrio cholerae*. *Science*, **310**, 1824–1827.

100. Gerisch, G; Lüderitz, O; Ruschmann,E. (1967) Antikörper fördern die Phagozytose von Bakterien durch Amöben. *Zsch Naturforsch.*, **22b**, 109.

101. de Souza Silva,O. and Blokesch,M. (2010) Genetic manipulation of *Vibrio cholerae* by combining natural transformation with FLP recombination. *Plasmid*, **64**, 186–195.

102. Blokesch,M. (2012) TransFLP - A method to genetically modify *Vibrio cholerae* based on natural transformation and FLP-recombination. *J. Vis. Exp.*, 10.3791/3761.

103. Borgeaud,S. and Blokesch,M. (2013) Overexpression of the *tcp* gene cluster using the T7 RNA polymerase/promoter system and natural transformation-mediated genetic engineering of *Vibrio cholerae*. *PLoS One*, **8**, e53952.

104. Dalia,A.B., McDonough,E.K. and Camilli,A. (2014) Multiplex genome editing by natural transformation. *Proc. Natl. Acad. Sci. U. S. A.*, **111**, 8937–8942.

105. Benghezal,M., Fauvarque,M.O., Tournebize,R., Froquet,R., Marchetti,A., Bergeret,E., Lardy,B., Klein,G., Sansonetti,P., Charette,S.J., *et al.* (2006) Specific host genes required for the killing of *Klebsiella* bacteria by phagocytes. *Cell. Microbiol.*, **8**, 139–148.

106. Keene,O.N. (1995) The log transformation is special. *Stat. Med.*, **14**, 811–819.

107. Lo Scrudato,M. and Blokesch,M. (2012) The regulatory network of natural competence and transformation of *Vibrio cholerae*. *PLoS Genet.*, **8**, e1002778.

108. Sommer,D.D., Delcher,A.L., Salzberg,S.L. and Pop,M. (2007) Minimus: a fast, lightweight genome assembler. *BMC Bioinformatics*, **8**.

109. Seemann,T. (2014) Prokka: Rapid prokaryotic genome annotation. *Bioinformatics*, **30**, 2068–2069.

5 A SINGLE NUCLEOTIDE DRIVES TYPE VI SECRETION IN *VIBRIO CHOLERAE*

Drawing by Noémie Matthey, PhD

5.1 INTRODUCTION

The main part of this introduction is an excerpt from the introduction of this thesis (chapter 1).

Vibrio cholerae, a comma-shaped Gram-negative bacterium and the causative agent of the disease cholera, belongs to the class of Gammaproteobacteria. Vibrios are abundant in marine environments and members of the genus *Vibrio* can be found associated with a wide diversity of organisms, such as sponges, corals, phyto- and zooplankton, fish, shellfish, etc. While some species are symbionts, many of them are important pathogens for marine organisms or humans, causing vibriosis and gastrointestinal diseases, respectively (1, 2).

Bacteria like *V. cholerae* face multiple selective pressures in the aquatic environment, such as inadequate access to nutrients, grazing pressure from predators and competition with other bacteria. To deal with such competition, bacteria invest a lot of their genomic space and energy into an arsenal of molecular weapons (3–5). Among these different bacterial molecular weapons, the type VI secretion system (T6SS) is of particular interest, since it is the only clear mechanism for bacterial antagonism found in *V. cholerae*, at least in well-studied pandemic strains (3). The T6SS is composed of many proteins that form a baseplate along with the VgrG/PAAR spike proteins, to which an inner tube (formed by Hcp hexameric rings) and a contractile VipA/B sheath that surrounds the Hcp tube are attached. Contraction of the sheath pushes the Hcp tube along with the VgrG/PAAR tip to the extracellular milieu or across a target cell's envelope (6, 7). The ClpV ATPase subsequently unfolds the contracted VipA/B sheath and recycles the subunits for a new assembly round of sheath extension and contraction. Importantly, effector toxins are secreted by the T6SS either as cargos bound to structural proteins or as C-terminal domains of PAAR, Hcp or VgrG (8). These effectors target conserved cellular components such as nucleic acids, membranes, peptidoglycan or the eukaryotic cytoskeleton (9, 10). Cells that produce the

T6SS protect themselves from the toxic activity of these effectors by expressing cognate immunity proteins, which are usually encoded next to the effector genes, forming effector/ immunity (E/I) pairs (11–14). Besides self-protection, this system is employed as a mode of kin discrimination, as bacteria producing the same effector and immunity proteins are able to coexist (14).

The T6SS of pandemic *V. cholerae* is encoded by four gene clusters: the large cluster and three auxiliary clusters (Aux 1-3) (15). The large cluster contains mainly genes encoding structural components of the machinery. It also harbors a gene encoding one of the effectors, namely an evolved VgrG3 protein with a C-terminal domain with lysozyme activity, and the corresponding immunity protein TsiV3 (12, 16). The large cluster also contains an activator of RpoN (σ54) called VasH, which induces the expression of Aux 1 and 2 (17). These auxiliary clusters are similar in architecture, each one with a copy of the genes encoding Hcp and VgrG. In addition, both Aux1 and 2 carry a gene encoding an adaptor protein (Tap1 or VasW (18, 19)) and an E/I pair, namely the bifunctional lipase TseL and immunity TsiV1 in Aux1 (13, 16, 20) and the pore-forming toxin VasX and immunity TsiV2 in Aux2 (10, 21, 22). Importantly, the large and Aux clusters 1 and 2 are conserved among *V. cholerae* strains. Aux3, on the other hand, is conserved in pandemic *V. cholerae* strains, while a mobile and prophage-like form of this cluster can be sporadically found in environmental isolates (23). Aux3 contains a second copy of PAAR (the first copy is found in the large cluster) (24) and the E/I pair TseH/I (25, 26).

The T6SS of pandemic *V. cholerae* strains is tightly regulated and kept silent under laboratory conditions, though it can be activated by two independent signaling pathways. *V. cholerae* is commonly found associated to chitin surfaces, which induce the physiological state of natural competence for transformation in this bacterium. This induction happens when cells reach high cell density (HCD) while growing on chitin and leads to the production

of the competence activators TfoX and QstR (27–29). Importantly, the T6SS is concomitantly produced under these conditions, therefore allowing the uptake of DNA released from killed non-kin neighboring bacteria (30). However, *V. cholerae* also contains another TfoX-like protein called TfoY (31), which, while having no role in natural competence induction, also induces the T6SS in pandemic *V. cholerae* (17). TfoY is produced as a response to low intracellular c-di-GMP levels, which is measured post transcriptionally by a riboswitch that is located at the 5'-end of the *tfoY* mRNA (17, 32). In sharp contrast to what is seen in pandemic strains, the diverse and well-distributed environmental isolates, as well as non-pandemic toxigenic *V. cholerae* strains, keep their T6SS constantly active (15, 33–37). The mechanism behind this constitutive T6SS production and activity in *V. cholerae* is currently not understood, and represents the focus of this study.

Here, we constructed a library of 800 hybrid clones using a transformation-based cross of a pandemic and a non-pandemic strain. These mosaic genome-carrying strains were then screened for their T6SS activity, which allowed us to find a causative single nucleotide polymorphism (SNP) that controls T6SS activity in *V. cholerae*. By studying the T6SS regulation linked to the identified SNP (nt45), we developed a preliminary working model describing the constitutive versus inducible T6SS production in environmental/non-pandemic and pandemic *V. cholerae* strains, respectively.

5.2 RESULTS AND DISCUSSION

5.2.1 Screening of a hybrid strain library for changed T6SS activity

We were interested in understanding how non-pandemic *V. cholerae* strains are able to keep their T6SS constitutively active, while pandemic strains only induce the machinery under specific conditions. To address this question, we first established a library of hybrid

strains where each clone contained a mosaic genome derived from the pandemic strain A1552 and the non-pandemic strain ATCC25872. The library design was based on our previous work on horizontal gene transfer between *V. cholerae* cells driven by natural transformation (38). To start with, we genetically engineered the T6SS-silent pandemic strain A1552 (17, 30, 39) by inserting antibiotic marker genes into its genome, resulting in the DNA donor strains. Specifically, we integrated the kanamycin-resistance gene *aph* at 40 different positions inside strain A1552's genome, whereby insertion was approximately 100-kb apart from each other in the individual strains. 30 *aph* marker genes were inserted on the 3-Mb large chromosome and 10 *aph* genes in the 1-Mb small chromosome (Figure 5.1a). Next, we isolated the genomic DNA (gDNA) of these 40 donor strains and added it individually to the competence-induced recipient strain ATCC25872 (40–42). After 30 h of co-incubation, 20 kanamycin-resistant hybrid clones were isolated from each of the 40 transformation experiments (Figure 5.1a). The rationale behind this library design was that large regions of DNA surrounding the resistance marker are co-exchanged with *aph* during transformation. Indeed, during transformation, *V. cholerae* cells acquire and integrate, on average, ~50 kb in length (including the resistance marker), while in 10-15% of transformants the exchanged region exceeds 100 kb (38). We therefore expected that one or two out of each set of 20 clones should have co-exchanged more than 100 kb with *aph*, thereby covering the entire genome using the 40 input gDNAs.

Next, these 800 hybrid clones were screened for their T6SS activity in an *E. coli* killing experiment, which we adapted based on a fluorescence imaging approach (inspired by the BaCoF method previously established by Ben-Yaakov and Salomon (43)). In these experiments, prey survival was scored by their GFP signal. Bacterial spots where prey was mixed with T6SS- predators would remain alive and therefore GFP+, while spots where the prey was killed by T6SS+ predators would result in a lack of a GFP signal. In the case of the

hybrid strains library, the original receiver strain was T6SS+. In our screen, we therefore scored hybrid clones that had lost T6SS activity when an *aph* gene and the adjacent genomic area were transferred (Figure 5.1a).

5.2.2 Transfer of *aph*#32 changes T6SS activity

Strikingly, 19 out of the 20 hybrid clones that had transferred the *aph* marker at position #32 from the pandemic strain lost their T6SS activity in our fluorescence-based screening (Figure 5.1b). To confirm these results, we created an inverse *aph*#32 library in which the *aph* marker was inserted into the same genomic position in the T6SS+ strain ATCC25872. The experiment was then repeated with the pandemic strain as competent acceptor strain, in which case 7 out of 20 transformants had switched T6SS activity (Figure 5.1b). The location of *aph*#32 was ~15 kb upstream of the large T6SS cluster (Figure 5.1c), suggesting that the region responsible for the T6SS activation might lie inside the large cluster itself. We therefore engineered an additional construct (#42) in which case the *aph* marker was integrated immediately upstream of the first gene of the large cluster (*paar1*; Figure 5.1c). Repeating the above-described transformation experiment using the gDNA of this new strain resulted in 20 out of 20 transformants that had a changed T6SS activity status (from + to -). Transfer of the resistance marker in the other direction, i.e., from the non-pandemic to the pandemic strain, resulted in 19 out 20 clones with a change in T6SS status (from – to +) (Supplementary figure 5.1a). We also added an additional *cat* marker (for chloramphenicol resistance) at the very end of the large cluster in combination with *aph*#42 to ensure transfer of the full large cluster. When both markers were integrated in the pandemic strain (T6SS-) and then transferred together to the non-pandemic strain (T6SS+), all 20 clones became T6SS- (Supplementary figure 5.1b), suggesting that the T6SS large cluster contains indeed the regulatory region.

5.2.3 A single nucleotide drives T6SS activity in *V. cholerae*

Next, we Sanger-sequenced selected regions from the T6SS large cluster to check whether they were transferred in these hybrid clones. As previously mentioned, transfer of *aph*#42 from the non-pandemic to the pandemic strain resulted in 19 out of 20 clones with a T6SS that was now active. The non-converted hybrid clone therefore served as our negative control. We initiated our survey by sequencing genes encoding the two known regulators found in the T6SS large cluster: *vasH* and *VCA0122*. Importantly, the two strains used in the library construction display sequence polymorphisms in these genes. VasH is an activator of RpoN and induces the expression of T6SS auxiliary clusters 1 and 2 (16, 17, 33, 44), while VCA0122 was proposed to control *hcp2* expression at the transcriptional level (16). However, we observed that many of the hybrid clones with a changed T6SS status had neither transferred *vasH* nor *VCA0122* (data not shown).

Next, we checked for polymorphisms between the two strains that were located towards the beginning of the large cluster, closer to where *aph*#42 was inserted. While doing so, we identified a distinctive single nucleotide polymorphism (SNP). This SNP (G in pandemic strain A1552 and T in non-pandemic strain ATCC25872) was located in a 431-bp intergenic region between the second (*VCA0106*) and third (*vipA*) genes of the large cluster, 44 bp downstream of *VCA0106*. When we Sanger-sequenced this region in the T6SS-converted hybrid transformants, we observed that they had acquired the SNP (from here onwards referred to as intergenic nucleotide 45 or “nt45”) from the non-pandemic strain, while the single non-T6SS-converted clone had retained the original nucleotide (Figure 5.1d). Next, we checked the *VCA0106-vipA* intergenic region in a collection of fifteen environmental *V. cholerae* isolates that also keep the T6SS constitutively active (45). Strikingly, all strains had a T nucleotide in the nt45 position (Figure 5.1e).

In order to check if nt45 was required and sufficient to determine T6SS activity, we performed site-directed mutagenesis. Specifically, we change the G to T in the pandemic strain and the T to G in the non-pandemic strain. As can be observed in Figure 5.1f, this nucleotide conversion was indeed sufficient to change the strains' T6SS activity. These results were further validated through T6SS visualization by fluorescence microscopy in which case the strains also carried a translation fusion of the T6SS sheath protein VipA and superfolder GFP (sfGFP) (Figure 5.1g).

Next, we exchanged nt45 in the pandemic strain background to either a cytosine (C) or an adenine (A), which are polymorphisms that we never detected in any of the thousands of *V. cholerae* genome sequences that are deposited in NCBI. Interestingly, the pyrimidine C also drove activation of the T6SS, even though this phenotype was less pronounced than with the T variant. The purine base A kept the T6SS silent alike the G variant in pandemic strains. Deletion of 10 bp around (and including) nt45 also abolished T6SS activity (Figure 5.1h).

Further comparative genomics based on the NCBI database highlighted that the T (T6SS+) variant of nt45 was well-conserved among non-pandemic *V. cholerae* strains, while the G variant was restricted to but fully conserved in the 7th pandemic lineage. To confirm that the nt45-linked effect on T6SS activity was conserved in other *V. cholerae* strains, we performed site-direct mutagenesis to change the nt45 in a panel of pandemic and environmental *V. cholerae* strains. Indeed, mutation of nt45 alone was sufficient to either activate the T6SS in a collection of 7th pandemic strains (Figure 5.2a) or to silence the T6SS in a set of environmental isolates (Figure 5.2b).

5.2.4 Nt45-conversion effects in T6SS transcript and protein levels

In order to understand the nature of T6SS regulation via nt45, we investigated how it affected T6SS transcript and protein levels. As can be seen in Figure 5.2c, the nucleotide conversion alone was sufficient to increase the transcript levels of T6SS genes in the pandemic strain and to decrease them in the non-pandemic strain background, respectively. A similar effect was also observed in other pandemic and non-pandemic isolates (Supplementary figure 5.1c-d). Importantly, the transcriptional effect was seen in genes belonging to the large and auxiliary clusters. Precisely, the sheath proteins-encoding genes *vipA/B*, which are located adjacent to the intergenic region, were affected, as was the downstream gene *vasH*, which is located towards the end of the large cluster. As expected from the effect on *vasH* expression and the requirement of the VasH regulator for Aux 1 and 2 expression, nt45 exchange also affected the transcript levels of both *hcp* copies (Figure 5.2c; Supplementary figure 5.1c-d).

Considering that we observed a nt45-related effect on T6SS transcript levels (by affecting transcription or mRNA stability), we next tested whether the expression effects extended beyond the T6SS clusters. Thus, we performed RNA sequencing (RNAseq) of the WT and nt45-converted variants of a pandemic (A1552) and a non-pandemic (ATCC25872) strain. As can be observed in Supplementary figure 5.2, the nt45-related expression change was narrow and mostly restricted to T6SS-related genes. Moreover, nt45 conversion affected only the transcript levels of genes downstream of its location, while expression of the two genes upstream, *paar1* and *VCA0106*, was unaffected (Supplementary figure 5.3; Supplementary tables 5.2 and 5.3).

Next, we tested if the observed nt45-conversion response in expression profile would translate into protein levels. We observed that nt45-conversion increased or decreased the amounts of the inner tube protein Hcp being produced, depending on the strain background

(Figure 5.2d). We also observed an effect on Hcp secretion (Figure 5.2d), which was consistent with *E. coli* predation data (Figure 5.1f). The effects on Hcp production and secretion were also validated in our panel of pandemic and environmental strains (Supplementary figure 5.1e-f).

5.2.5 HapR regulates T6SS secretion at high cell density

The T6SS of pandemic *V. cholerae* is kept silent under laboratory conditions, but it can be activated by at least two independent routes. One involves the master and intermediate regulators TfoX and QstR, respectively, which induce both the T6SS and the DNA uptake machinery when cells grow to HCD on chitin (27–30). Another route works independently of chitin or quorum sensing (QS) and involves TfoY, which activates the T6SS once intracellular c-di-GMP levels decrease (17). A putative environmental signal that might induce this low c-di-GMP state has yet to be identified. In this study, we wondered whether any of these three regulators were somehow involved in the T6SS control via nt45. We therefore constructed strains that either lacked any of these three regulators or the master regulator of QS in *V. cholerae*, HapR. These strains were used as predators in *E. coli* killing assays. While a lack of T6SS structural proteins (VipB or VasK) completely abrogated the T6SS's activity in the nt45-converted pandemic strain, none of the known T6SS regulators TfoX, QstR or TfoY played a role in this context (Figure 5.3a). These results were confirmed by qRT-PCR and Western blot analyses, where deletion of these regulators did neither impact T6SS transcript levels nor Hcp production and secretion (Supplementary figure 5.4). When HapR was deleted in the nt45-converted pandemic strain, a mild effect on *E. coli* killing was observed (Figure 5.3a). We therefore profiled the transcripts of WT and Δ*hapR* variants in cells grown for 3 h (mid-exponential phase) or 6 h (late exponential phase) and observed that *hapR*-minus strains showed slightly lower T6SS transcript levels,

independently of the growth phase (Figure 5.3b). Next, we checked the effect of the presence or absence of HapR on the Hcp protein level in the nt45-converted strain. At 3 h of growth, *hapR* deletion caused a slight decrease in Hcp protein production and secretion (Figure 5.3c), in accordance with the transcript levels. Surprisingly, at 6 h of growth we observed decreased Hcp secretion for the parental strain, whereas the *hapR* mutant kept secreting Hcp abundantly. Our results therefore suggest that HapR represses T6SS-mediated secretion at HCD, as Hcp production itself was not significantly changed (Figure 5.3d). We confirmed these results by quantifying VipA-sfGFP (T6SS) structures/ cell in the parent and Δ*hapR* variants of the T6SS-active pandemic strain. At 3 h, the parental strain displayed ~1.4 T6SS structures/ cell, which decreased to ~0.35 structures/ cell at HCD. The Δ*hapR* variant contained ~0.85 and ~0.6 T6SS structures/ cell at 3 h and 6 h, respectively (Figure 5.3e).

Even though the exact mechanism behind this HapR-regulation is still not clear, we obtained preliminary evidence that it might involve the HapR-activated hemagglutinin/ protease (H/A protease or HapA). HapA is an extracellular zinc metalloprotease that degrades several proteins, including host's mucins (46). Notably, in our experiments a Δ*hapA* strain continued secreting Hcp at HCD comparable to the Δ*hapR* strain (Figure 5.3d). We therefore hypothesize that the effect exerted by HapR on Hcp secretion is dependent on the HapR-dependent activation of *hapA*. This line of evidence needs further investigation, especially given the fact that HapA is secreted by the type II secretion system (47, 48). Interestingly, a recent study evaluated abundance of T6SS proteins by targeted proteomics in different bacteria with an emphasis on *V. cholerae* (49). The authors observed a steady decrease in the abundance of T6SS proteins while cell progressed from early exponential phase towards later growth phases, which was also quantified by the amount of T6SS structures/ cell. The instability of two proteins in particular, TssE (HsiF) and VasX, seemed

to explain this decrease in T6SS assembly, even though the mechanism by which these proteins were degraded remained unanswered (49).

5.2.6 Promoter identification in nt45-containing intergenic region

The nt45-containing intergenic region is considerably long (431 bp); thus, we were wondering which part of it was essential for T6SS function. We therefore engineered strains in which portions of the intergenic region were gradually removed. For all constructs, we kept 10 bp downstream of *VCA0106* and 25 bp upstream of *vipA*, to preserve any important regulatory regions, like the ribosomal binding site of *vipA*. We removed either ~300 or ~350 bp from the intergenic region, keeping the nt45 as either a G (T6SS-) or a T (T6SS+). We also generated a strain lacking the entire intergenic region apart from the 10/ 25 bp mentioned above (Figure 5.4a). Next, we tested these strains in *E. coli* killing assays (Figure 5.4b). Removal of the entire intergenic region abrogated T6SS activity, indicating that no repression of T6SS activity occurs by this region in the pandemic strain, for instance by the insertion of a transcriptional terminator (Figure 5.4b). On the other hand, in the case of the nt45-converted strain, the Δ300bp strain kept the same level of T6SS activity as its parental strain, while the Δ350bp strain displayed no T6SS function anymore. This result suggested that the region lacking in the Δ350bp strain but not in the Δ300bp strain had an important regulatory function (Figure 5.4c). We then engineered a strain lacking only this region (which is ~60bp long), and it showed a complete loss of T6SS activity (Figure 5.4b), supporting our notion on the necessity of this ~60bp region.

Next, we examined this ~60bp region and noticed a possible -10/ -35 box promoter sequence (Figure 5.4c), which we mutated accordingly in the nt45-coverted strain (e.g., the -10 box TAG<u>AA</u>T to TAG<u>GC</u>T). This mutation once again abolished T6SS activity (Figure

5.4d). The same mutation was also introduced into the nt45-non-converted WT pandemic strain to test its effect upon artificial TfoX, QstR, or TfoY induction (e.g., in strains carrying arabinose-inducible copies of *tfoX*, *qstR* or *tfoY)*. Interestingly, the changed -10 box also impacted T6SS induction under these conditions (Figure 5.4d), indicating that all regulatory circuits involved this potential promoter region to induce T6SS production. Finally, we observed a slight nt45-dependent response in transcript levels of the intergenic region right after the putative promoter by qRT-PCR (Figure 5.4e).

5.2.7 Nt45 activity most likely involves a regulator encoded in the large cluster

In order to gain further insight into the dynamics of T6SS regulation via nt45, we established an approach using the VipA-sfGFP translational fusion as a reporter for T6SS activity under different settings. The fusion-encoding gene was either inserted at *vipA*'s native chromosomal locus close to the start of the large T6SS cluster or inserted into a pBAD-Kan plasmid (50) for expression *in trans*. In the latter case, the *vipA-sfGFP* allele was proceeded by a 3-kb region that included a ~1kb region upstream of *paar1*, *paar1* & *VCA0106*, and the intergenic region (containing nt45 as G or T).

When these two reporter plasmids were electroporated into a WT pandemic strain, we only observed VipA-sfGFP production for the "T" version of nt45 but not for the "G" version (Figure 5.5a). On the other hand, when the plasmids were electroporated into the nt45-converted pandemic strain (T6SS+), VipA-sfGFP expression occurred from both plasmids (Figure 5.5b). This result suggested that the nt45-driven T6SS activation could happen *in trans* when the T6SS-positive nt45 was located on the chromosome but the *vipA-sfGFP* reporter (even with a nt45-G) was on the plasmid. We therefore wondered whether the same could happen in the opposite direction, i.e., if the nt45-converted region located

on the plasmid could activate a chromosomal *vipA-sfGFP* reporter. Interestingly, this setting did not induce the production of VipA-sfGFP (Figure 5.5c). This result suggested to us that the T6SS induction might require an additional regulator that is located downstream of *vipA* and therefore inside the large cluster. This region would be lacking in the plasmid-carried construct, as it did not extend beyond the *vipA-sfGFP* gene.

As mentioned previously, the T6SS large cluster in *V. cholerae* encodes two known regulators: VasH and VCA0122 (16, 17, 33, 44) and we therefore wondered whether they could be involved in the here-studied regulatory circuit. Accordingly, we restored the condition in which we observed *in trans* T6SS activation (with nt45-T on chromosome and the *vipA-sfGFP* allele with nt45-G on the plasmid) and tested the impact of deleting *vasH* or *VCA0122*. As shown in figure 5.5d, however, the fusion was still activated under these conditions, demonstrating that VasH and VCA0122 are dispensable for the nt45-mediated T6SS regulation (apart from VasH's requirement for auxiliary clusters 1 and 2 activation).

Finally, we decided to engineer a strain lacking the majority of the T6SS large cluster. This strain still contained *paar1* & *VCA0106*, the intergenic region (with the "T" nt45), the beginning of *vipA* and a fragment of the immunity gene *tsiV3* at the end of the cluster. When this strain that also carried the nt45-G *vipA-sfGFP* reporter plasmid was checked by microscopy, we observed a loss of VipA-sfGFP production (Figure 5.5d). This result supports our notion that the large cluster encodes a so-far unidentified regulator that generates a positive feedback loop to fully induce T6SS production (Figure 5.6).

5.3 CONCLUSION AND FUTURE PERSPECTIVES

While the T6SS regulation is well-studied in pandemic *V. cholerae* strains (17, 30, 51), the more widespread constitutive T6SS activity in non-pandemic *V. cholerae* has puzzled the

field since the T6SS discovery in 2006. Through an innovative approach using a library of hybrid strains, our study uncovered a single nucleotide that causes constant T6SS production and activity. Importantly, this constitutive T6SS activity seems to be ancestral, while the controlled T6SS activity has evolved in the 7th pandemic lineage. Notably, a pre-7th pandemic El Tor strain, MAK757, collected in Celebes Islands in 1937, still carries the activating "T" as nt45, consistent with its active T6SS (data not shown). Furthermore, T6SS transcript and protein levels are further elevated in the T6SS-active non-pandemic strains upon artificial induction by TfoX, TfoY or QstR (data not shown), demonstrating that these regulatory circuits were kept, rather than evolved, in the 7th pandemic lineage. It is tempting to speculate that having "on demand" T6SS might contribute to the success of the 7th pandemic lineage in the disease context. Conversely, keeping the machinery constantly active might be useful for *V. cholerae* living in the aquatic environment, where defense against predators and bacterial competitors might be constantly required.

Our results also highlighted that pandemic strains show low background T6SS activity, which is completely abrogated if the nt45 region is deleted (Figure 5.1h). We discovered that this activity was based on random T6SS activation in approximately one in fifty cells, as we showed by microscopy (Figure 5.1g; counting data not shown). This fact combined with the observations described above, suggest that the nt45-dependent regulation probably requires a positive feedback loop (Figure 5.6). Indeed, in our model, the nt45-converted condition, which resemble the non-pandemic strains, allows transcription and translation of the large cluster and a regulatory gene within this region. The encoded regulator might then feedback to the putative promoter to further induce the region's transcription. On the other hand, in the nt45-G situation in the native pandemic strains, the initial expression of the T6SS cluster and the regulatory gene might be low and only rarely

(in 1 out of 50 cells) exceed the transcriptional noise level to trigger the positive feedback loop.

The final step of this project will be to identify the regulator gene (encoding a protein, small RNA, peptide...) in the large cluster. To do so, we will engineer strains with truncated versions of the large cluster and test them with our reporter plasmid. Even though one of the large cluster proteins with known function might play an additional but so far unstudied role in the here-described nt45-dependent regulation, the cluster also contains several genes with unknown roles. For instance, for the gene product of *vasI* (also known as *tagO*), BlastP analysis does not indicate any known domains, and the gene is not widely present outside the *Vibrio* genus in T6SS loci of other bacterial species (52). VasI was shown be dispensable for Hcp production and secretion in the non-pandemic strain V52 and barely affected its *E. coli* killing capability. On the other hand, VasI was necessary for amoebal intoxication, apparently because it contributes to the actin-crosslinking activity of VgrG1 (16).

Finally, *fha* (or *vasC*) was demonstrated to be essential for T6SS assembly (53). The intracellular Fha protein level is high in the non-pandemic *V. cholerae* strain 2740-80 (49), but its exact function is still unknown. Fha is a FHA (forkhead-associated)-domain containing protein, and members of this family are encoded in T6SS loci from different bacterial species (52). Indeed, Fha is essential for the tit-for-tat T6SS regulation in *Pseudomonas aeruginosa* (54). In this organism, sensing of a signal by the periplasmic protein TagR leads to the dimerization and autophosphorylation of the transmembrane serine-threonine kinase PpkA, followed by the phosphorylation of Fha, which ultimately promotes T6SS firing (55–57). *fha* transcripts were abundant and responsive to the nt45 conversion in our RNA-seq experiments (Supplementary tables 5.2 and 5.3). That fact, in addition to its clear regulatory role in *P. aeruginosa,* suggests that Fha could be involved in the nt45-mediated regulation.

5.4 MATERIAL AND METHODS

5.4.1 Bacterial strains and growth conditions

Bacterial strains used in this study are listed in Supplementary table 5.1. Unless otherwise stated, strains were grown aerobically in Lysogeny broth (LB; 10 g/ L of tryptone, 5 g/ L of yeast extract, 10 g/ L of sodium chloride; Carl Roth) or on LB agar plates at 30 °C or 37 °C. Half-concentrated defined artificial seawater medium (0.5×DASW) containing HEPES (4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid; Sigma-Aldrich) and vitamins (27) or 0.5X HW Marine Mix (Wiegandt, Germany) were used for growth on chitin (see below for natural transformation-based strain construction). Thiosulfate citrate bile salts sucrose (TCBS, Sigma-Aldrich) agar was used to counter-select *E. coli* following bacterial mating. Counter-selection based on SacB was performed on NaCl-free medium containing 10% sucrose. When required, the following antibiotics were added at their given concentrations: kanamycin (75 µg/ mL), chloramphenicol (2.5 µg/ mL), streptomycin (100 µg/ mL) and ampicillin (100 µg/ mL). To induce expression from the P_{BAD} promoter (from Tn*tfoX*, Tn*qstR* and Tn*tfoY*), cultures were grown in media supplemented with 0.2% L-arabinose.

5.4.2 Recombinant DNA techniques and genetic engineering

DNA manipulations and *E. coli* cloning were carried out using standard methods. PCR amplifications were performed using GoTaq (Promega), Pwo (Roche) and Expand High Fidelity (Roche) polymerases according to the suppliers' recommendations. Genetically modified loci were checked by Sanger sequencing (Microsynth, Switzerland).

Genetic engineering of *V. cholerae* was done using a combination of different methods. Natural transformation on chitin flakes and FLP-recombination were used to integrate the *vipA-sfGFP* (34) translational fusion construct allele into the chromosome of

diverse strains. Natural transformation on chitin flakes was also used for the hybrid strain library construction (explained in detail below), and for deletion of *vasH*, *VCA0122* and *hapA*. All remaining mutations and deletions were performed using allelic exchange by bi-parental mating using the pGP704-Sac28 counter-selectable plasmid (58).

5.4.3 Construction and screening of a hybrid strain library

The kanamycin-resistance gene *aph* was integrated at 40 different positions in the genome of pandemic *V. cholerae* strain A1552 (each position in an independent strain, resulting in 40 donor strains). Genomic DNA (gDNA) of these 40 strains was purified from 2 ml of an overnight culture, and DNA extraction was performed using 100/G Genomic-tips together with a Genomic DNA buffer set as described in the manufacturer's instructions (Qiagen). The non-pandemic acceptor strain ATCC25872 was grown on chitin flakes (40 different tubes in parallel) and each donor strain-derived gDNA was added to a tube and incubated at 30 °C for 30 h in total. After incubation, bacteria were selected on LB plates containing kanamycin and 20 clones were isolated and stocked from each independent transformation experiment (800 in total).

From this initial library, the clones were screened for their T6SS activity using a fluorescence-based *E. coli* killing experiment, which was adapted from (43). Briefly, 200 µL of the predator hybrid clones were mixed with 40 µL of GFP-labelled *E. coli* prey on 96-well plates, and 5 µL of the mixtures were spotted onto LB plates. The original donor (A1552) and acceptor (ATCC25872) strains were also mixed with GFP-labelled *E. coli* and included in all experimental plates, as controls. After 4 h incubation at 37 °C, plates were observed under a stereo microscope equipped with a green fluorescence lamp. Prey survival was scored based on their GFP signal (GFP+: predator was T6SS-; GFP-: predator was T6SS+ and killed the prey).

5.4.4 *E. coli* killing assays

Bacterial killing was assessed following a previously established assay (30). *E. coli* prey and the respective predator were mixed at a ratio of 1:10 and spotted onto paper filters on prewarmed LB agar plates (with 0.2% arabinose where indicated). After 4 h of incubation at 37 °C, bacteria were resuspended, serially diluted, and spotted onto antibiotic-containing (streptomycin, kanamycin or chloramphenicol) LB agar plates to enumerate the colony-forming units (shown as CFU/ mL). Statistically significant differences were determined on log-transformed data (59) by pairwise two-tailed Student's t-test of biologically independent replicates. If no prey bacteria were recovered, the detection limit was used to perform the statistical analysis.

5.4.5 Microscopy

Cells were mounted on microscope slides coated with a thin agarose pad (1.2% in 0.5X PBS), covered with a coverslip, and observed using a Zeiss LSM 700 inverted microscope (Zeiss, Feldbach, Switzerland). Images were prepared for publication using Fiji (60). Number of T6SS structures/ cell was estimated manually (900 – 1100 randomly chosen cells were counted per sample and experiment) and two biologically independent experiments per condition.

5.4.6 Quantitative Reverse Transcription PCR (qRT-PCR)

To analyze gene expression using quantitative reverse transcription PCR (qRT-PCR), overnight cultures were back-diluted 1:100 in LB medium and grown for 3 h or 6 h at 30 °C with agitation. RNA purification, DNase treatment, cDNA synthesis and qPCR followed a previously established protocol (61). Samples were analyzed on a LightCycler Real-Time

PCR System (Roche) using the standard curve method. Transcript levels in graphs are presented relative to the mRNA levels of the reference gene *gyrA* (gyrase).

5.4.7 RNA sequencing

Overnight cultures were back-diluted 1:100 in LB medium and grown for 3 h at 30 °C with agitation. Cells were harvested by centrifugation at 4 °C and washed with PBS buffer (Life Technologies), followed by lysis with Tri Reagent (Sigma-Aldrich) and shock freezing in a dry-ice ethanol bath. The samples were stored at -80 °C prior to further processing. RNA preparation and DNase treatment were performed as previously described (61). After DNase treatment, an additional purification step was performed using the GenElute Mammalian Total RNA miniprep kit (Sigma-Aldrich).

Downstream processing of the samples was conducted by Microsynth (Switzerland). Briefly, Illumina's TruSeq Stranded Total RNA Library Prep Gold kit including ribodepletion was used to construct libraries from total RNA. Subsequently, the Illumina NextSeq 550 platform and a high output v2.5 kit (single-end 1x75bp) were used to sequence the libraries. The produced single-end reads which passed Illumina's chastity filter were subject to de-multiplexing and trimming of Illumina adaptor residuals using Illumina's bcl2fastq software version v2.20.0.422 (no further refinement or selection). Quality of the reads in fastq format was checked with the software FastQC (version 0.11.8) (http://www.bioinformatics.babraham.ac.uk/projects/fastqc/). The reads were mapped to the reference genome via bowtie2 (version 2.3.5.1) (62) in local mapping mode with very sensitive pre-settings. To count the uniquely mapped reads to annotated genes, the software htseq-count (HTSeq version 0.11.2) (63) was used. Normalization of the raw counts and differential gene expression analysis was carried out with help of the R software

package DESeq2 (version 1.22.2) (64). Visualization of sample gene expression profiles was realized with the R software package Rtsne (version 0.15) (65).

Furthermore, mapped reads were provided as bam files and visualized using the Integrative Genomics Viewer (https://www.broadinstitute.org/igv/home). RNA-seq was performed for each of the four conditions (WT or nt45-converted samples for A1552 or ATCC25872) based on three independent biological replicates for each sample.

5.4.8 Western blotting

To check the production of Hcp and HapR proteins, cell lysates were prepared as described previously (17). In brief, overnight cultures were back-diluted 1:100 in LB medium and grown with agitation at 30 °C for 3 h or 6 h. Cells were harvested by centrifugation and the pellet was resuspended in 2X Laemmli buffer (Sigma-Aldrich), adjusting for the total number of bacteria according to the cultures' OD_{600} values. To check for T6SS-secreted Hcp, 1.5 ml of the culture supernatant was filter sterilized (0.2 µm filter; VWR) and the proteins were precipitated using trichloroacetic acid (TCA). The precipitated proteins were washed with acetone before being resuspended in 30 µl of 2X Laemmli buffer (Sigma-Aldrich). All samples were heated at 95 °C for 15 min.

Proteins were separated by sodium dodecyl sulfate (SDS) - polyacrylamide gel electrophoresis (PAGE) using 15% gels and then Western blotted as previously described (61). Primary antibodies against Hcp (Eurogentec; immunization no. ZCH15017) and HapR (Biomatic, immunization no. SA0033) were used at a 1:5,000 dilution. Loading controls were performed with anti-Sigma70-horseradish peroxidase (HRP) (BioLegend, cat. no. 663205), diluted 1:10,000, or anti-RNAP-beta (Neoclone; cat. no. WP001), diluted 1:5,000. Anti-rabbit IgG HRP (diluted at 1:20,000; Sigma-Aldrich, cat. no. A9169) was used as the secondary

antibody against the anti-Hcp and anti-HapR primary antibodies. Anti-mouse IgG HRP (diluted 1:20,000; Sigma-Aldrich, cat. no. A5278) was used as the secondary antibody against the anti-RNAP-beta primary antibody. Lumi-Light^{PLUS} western blotting substrate (Roche, Switzerland) served as the HRP substrate. The signals were detected using a ChemiDoc XRS+ station (BioRad).

5.4.9 Construction and electroporation of VipA-sfGFP reporter plasmids

VipA-sfGFP reporter plasmids were constructed in the pBAD-Kan (50) backbone and are described in Supplementary table 5.1. All cloned inserts were checked by Sanger sequencing (Microsynth, Switzerland). Plasmids were isolated using the GenElute HP Plasmid Miniprep kit (Sigma-Aldrich) and introduced into electrocompetent *V. cholerae* cells by electroporation at 1.6kV.

5.5 FIGURES

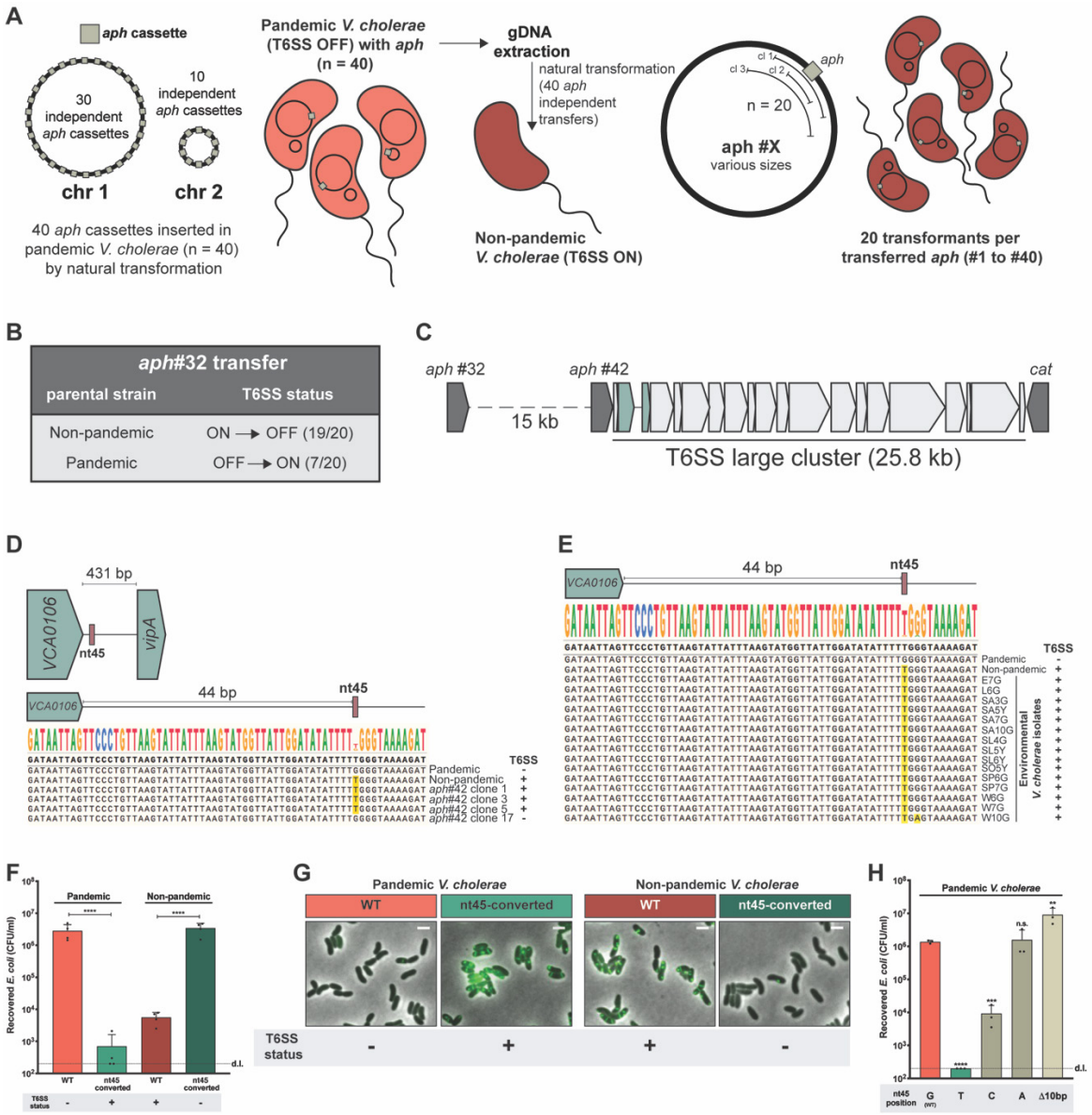


Figure 5.1 A single nucleotide, nt45, controls T6SS activity in *V. cholerae.*

(A) Schematic representation of the hybrid strain library construction. Details are given in the text. **(B)** T6SS activity scores of hybrid strains when the *aph*#32 resistance marker was transferred from a pandemic to a non-pandemic *V. cholerae* strain and vice-versa. **(C)** Scheme representing the T6SS large cluster of *V. cholerae* and the location of *aph*#32, *aph*#42 and *cat* markers. **(D)** Scheme representing the 431-bp intergenic region between genes *VCA0106* and *vipA* in the T6SS large cluster with nt45 located 44-bp downstream of *VCA0106*. Alignment shows part of the region in representative *aph*#42 transformants (with their T6SS status indicated on the right). **(E)** Alignment of the important part of the intergenic region in a collection of T6SS-active environmental *V. cholerae* isolates. **(F-H)** Site-directed mutagenesis leading to nt45 conversion results in changed T6SS phenotypes in pandemic and non-pandemic *V. cholerae*. **(F, H)** Bacterial killing assays using *E. coli* as prey. Numbers of surviving prey are depicted on the Y-axis (CFU/ml). Bar plots represent the average of at least three independent biological replicates (±SD). d.l., detection limit. Statistical significance is indicated (n.s., not significant; ** $P < 0.01$; *** $P < 0.001$; **** $P < 0.0001$). **(G)** Fluorescence microscopy of WT and nt45-converted variants of pandemic and non-pandemic *V. cholerae*. All strains produce a translational fusion of the T6SS sheath protein VipA and sfGFP. Phase contrast images were overlaid with the signal from the green fluorescent channel (GFP). Scale bars: 2 µm.

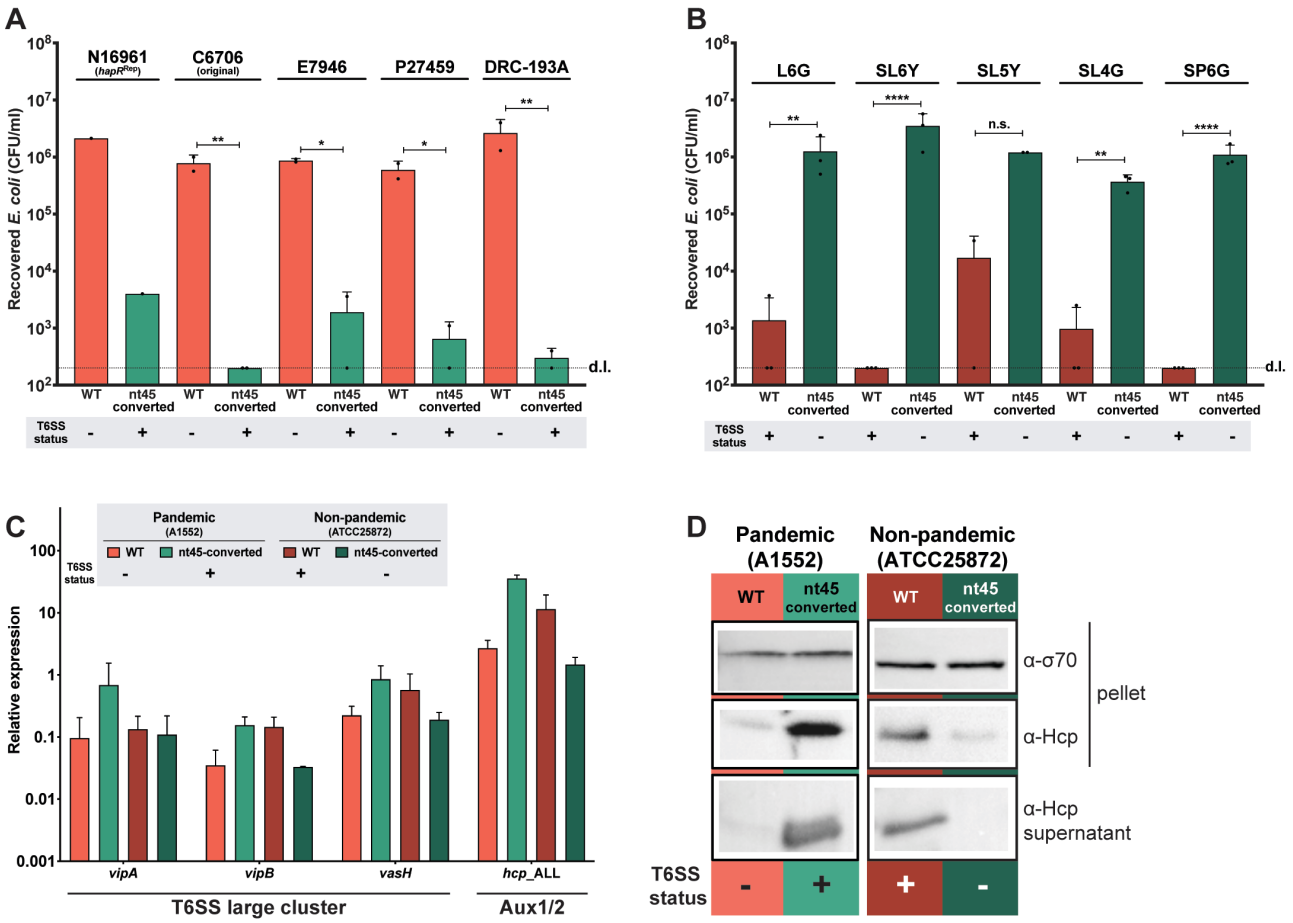


Figure 5.2 Nt45 conversion affects T6SS activity in diverse pandemic and environmental *V. cholerae* strains and leads to changes in T6SS RNA and protein levels.

(A, B) Site-directed mutagenesis was used for nt45 conversion in a collection of 7th pandemic **(A)** and environmental **(B)** strains, and their T6SS activity was assessed in *E. coli* killing assays. Numbers of surviving prey are depicted on the Y-axis (CFU/ml). Bar plots represent the average of at least two independent biological replicates (±SD), except for WT and nt45-converted N16961 strains, which were tested only once so far. d.l., detection limit. Statistical significance is indicated (n.s., not significant; * $P < 0.05$; ** $P < 0.01$; *** $P < 0.001$; **** $P < 0.0001$). **(C)** Expression scoring of T6SS genes for cells grown for 3h as determined by qRT-PCR in the indicated strains. Data represent the average of two independent biological replicates (±SD). **(D)** Hcp protein production and secretion in response to nt45 conversion in pandemic and non-pandemic *V. cholerae* after 3 h of growth. Intracellular (pellet) and secreted (supernatant) Hcp were assessed by immunoblotting using Hcp-directed antibodies. Detection of σ70 served as a loading control for the pellet fraction.

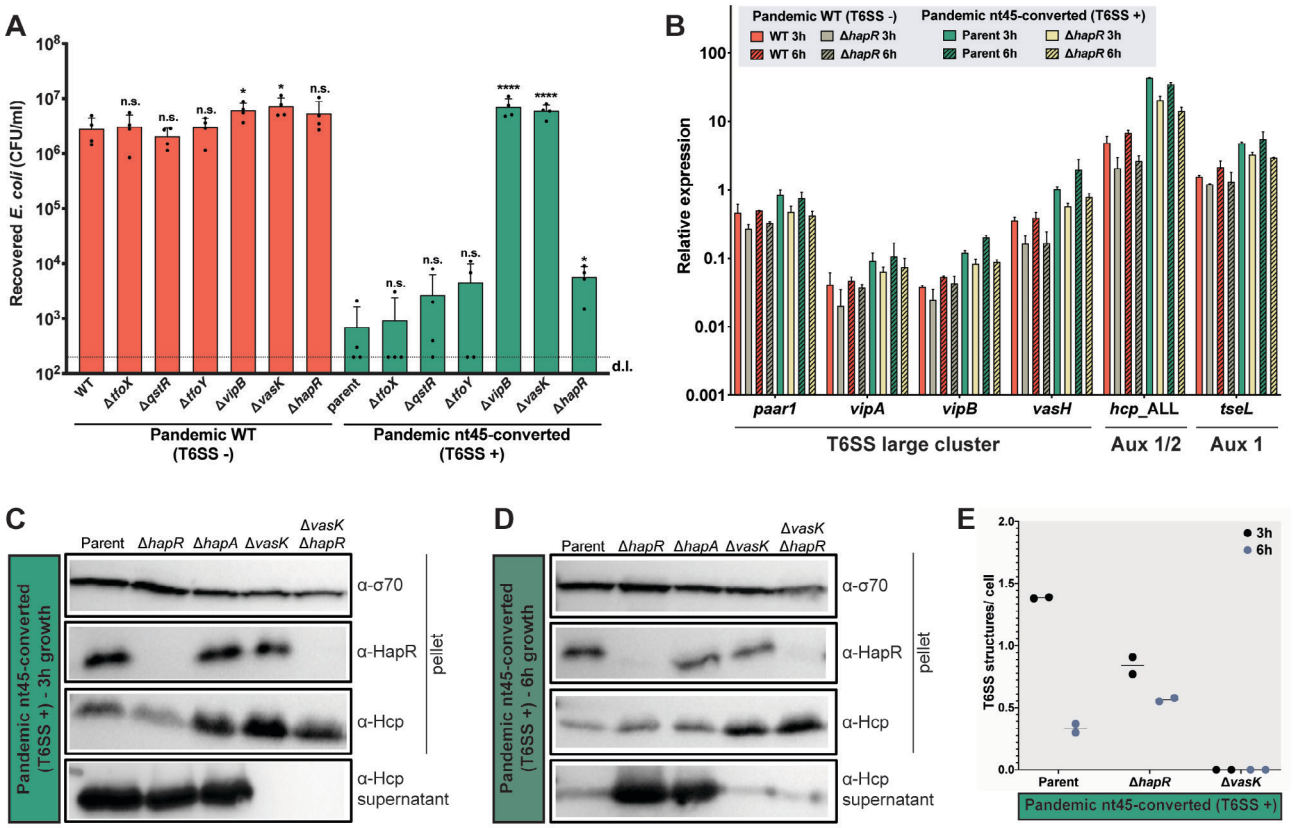


Figure 5.3 HapR controls T6SS secretion at high cell density.

(A) Strains lacking known T6SS regulators TfoX, QstR and TfoY, as well as T6SS genes *vipB* and *vasK* and the quorum-sensing master regulator HapR were constructed in the background of a WT pandemic strain (T6SS-) and its nt45-converted variant (T6SS+). These strains were tested as predators in bacterial killing assays using *E. coli* as prey. Numbers of surviving prey are depicted on the Y-axis (CFU/ml). Bar plots represent the average of four independent biological replicates (±SD). d.l., detection limit. Statistical significance is indicated for the deletion strains compared to their parentals (n.s., not significant; * $P < 0.05$; **** $P < 0.0001$). **(B)** Relative expression of T6SS genes was determined by qRT-PCR in WT and Δ*hapR* strains that were grown for 3 h or 6 h. Data represent the average of two independent biological replicates (±SD). **(C-D)** Strains constructed in a nt45-converted pandemic strain background (T6SS+) were harvested after being grown for 3 h **(C)** or 6 h **(D)** and intracellular (pellet) Hcp, HapR, and σ70 levels as well as secreted (supernatant) Hcp levels were assessed by immunoblotting using the respective antibodies. Note that HapR detection for samples grown to 3 h and 6 h was performed on different blots and the exposure times were chosen for best protein detection. Detection of σ70 served as a loading control for pellet fraction. **(E)** Number of T6SS structures/cell was scored by microscopy in different strains carrying a translational fusion of the T6SS sheath protein VipA-sfGFP, after growing the bacteria for 3 h or 6 h in LB. Between 900 and 1100 cells were counted per condition in each independent biological replicate. Each circle represents one replicate, and each line represents the average of these replicates.

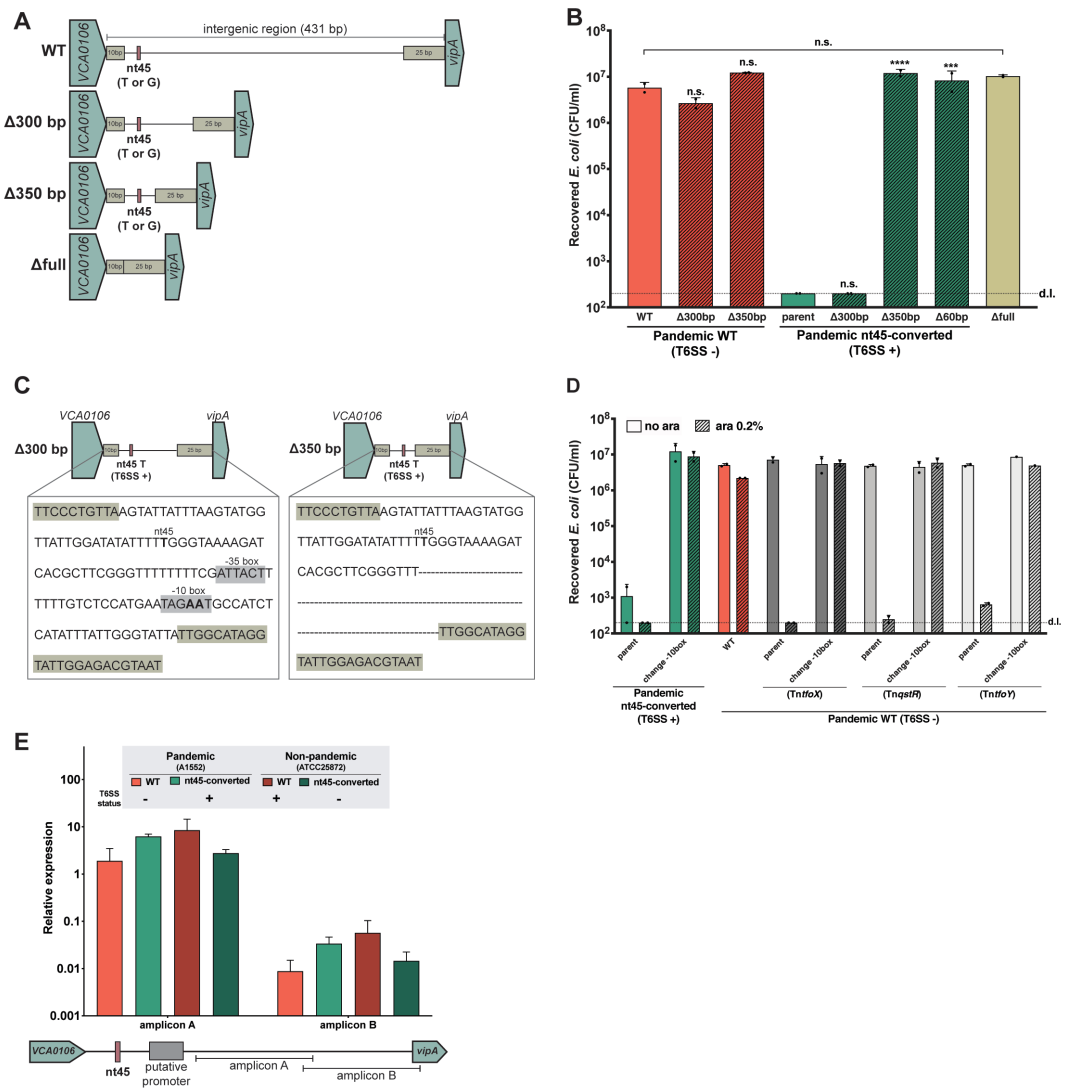


Figure 5.4 Characterization of nt45-containing intergenic region.

(A, B) The 431 bp-long intergenic region between *VCA0106* and *vipA* was genetically modified to assess which portions are required for promotion of T6SS activity. **(A)** Schemes represent genetic modifications performed. 10 bp downstream of *VCA0106* and 25 bp upstream of *vipA* were kept in all strains, but portions of the region in-between were gradually removed (Δ300bp; Δ350bp; Δfull). **(B)** T6SS activity of these strains (and also of a strain lacking only the 60bp region) was assessed by *E. coli* killing assays. Numbers of surviving prey are depicted on the Y-axis (CFU/ml). Bar plots represent the average of two independent biological replicates (±SD). Statistical significance is indicated for the deletion strains compared to their parentals (n.s., not significant; *** $P < 0.001$; **** $P < 0.0001$). **(C)** Representation of the nt45-containing intergenic region from the T6SS+ Δ300bp strain and the T6SS- Δ350bp strain. Note that the Δ350bp lacks the ~60-bp region which contains a predicted promoter (-10 and -35 boxes are showed in grey; the boldfaced AA nucleotides were changed to GC in the respective -10 box mutants). **(D)** Quantification of surviving *E. coli* prey after co-culture with variants of pandemic strains on LB agar without (no ara) or with (ara 0.2%) arabinose. Predator strains with an intact or mutated -10 box in the different backgrounds were tested: a nt45-converted strain (T6SS+) and the nt45-non-converted strain (T6SS-) without or with an arabinose-inducible copy of *tfoX* (Tn*tfoX*); *tfoY* (Tn*tfoY*) or *qstR* (Tn*qstR*). Numbers of surviving *E. coli* prey are depicted on the Y-axis (CFU/ml). Bar plots represent the average of two independent biological replicates (±SD). d.l., detection limit. **(E)** Relative expression levels of two amplicons amplified from the region that is immediately downstream of the putative promoter (A) or the region upstream of *vipA* (B). Details on strains is indicated in the legend. Bar plots represent the average of two independent biological replicates (±SD).

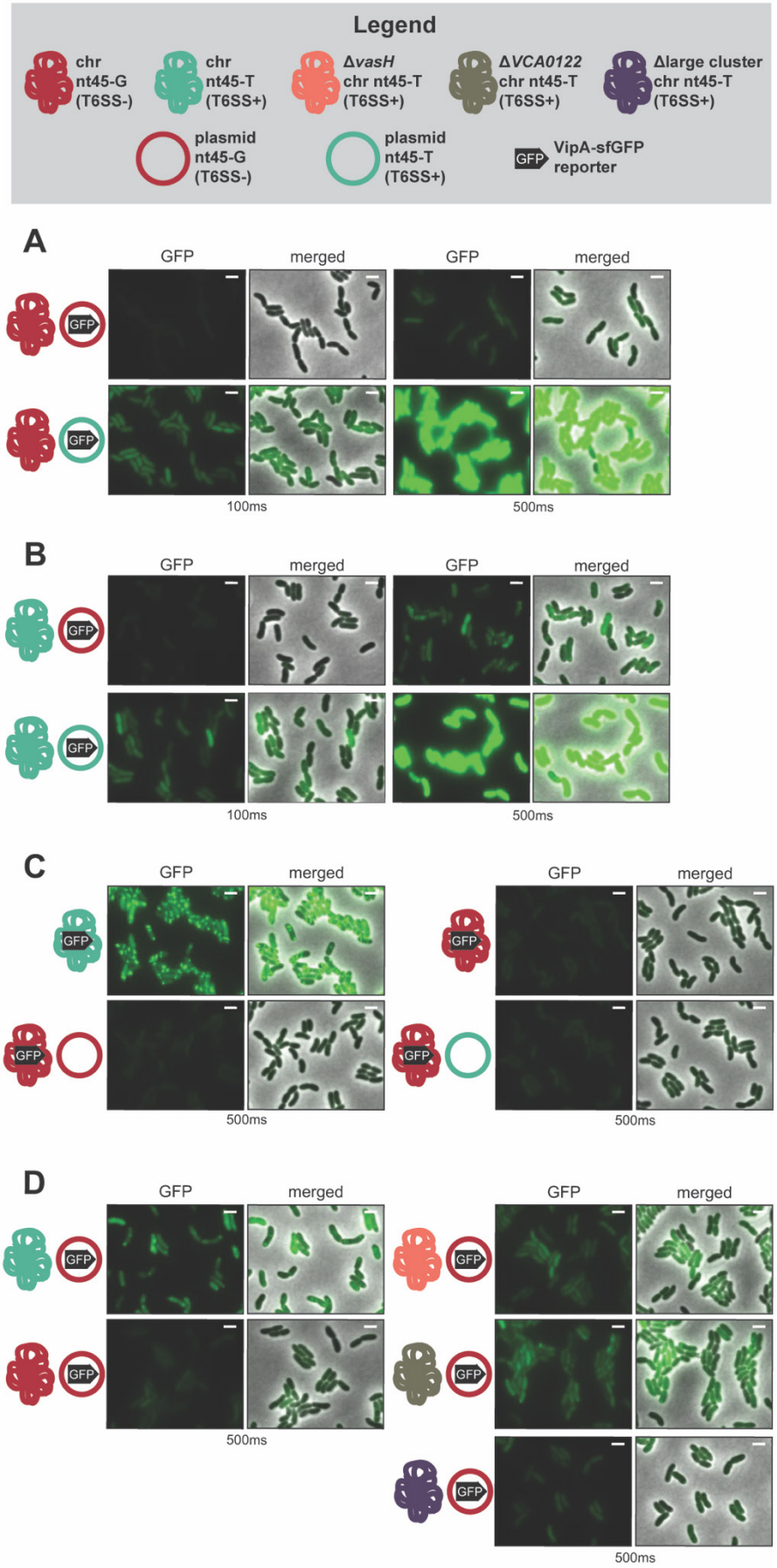


Figure 5.5 The T6SS large cluster likely contains a regulator that interacts with nt45 to control T6SS activity.

The VipA-sfGFP translational fusion was used as a reporter to assess T6SS expression. The fusion construct (black arrow) was inserted either at the native locus of *vipA* on chromosome 2 (shown inside packed DNA symbol) or in pBAD-Kan plasmid (shown inside circle). The chromosomal and plasmid nt45 are indicated by colors (red: G [T6SS-]; green: T [T6SS+]). **(A, B)** Pandemic *V. cholerae* strains with a nt45-G **(A)** or nt45-T **(B)** carrying VipA-sfGFP reporter plasmids with a nt45-G or T of their own. **(C)** Strains with nt45-T or nt45-G carrying the VipA-sfGFP fusion in the chromosome. In these cases, the plasmid, when present, contains only the region until *vipA* (without *sfGFP*). **(D)** Control conditions (left) are compared to strains (right) with a nt45-T in which the regulator-encoding genes *vasH* and *VCA0122* were deleted, as well as the T6SS large cluster (deletion starts in *vipA* and extends until the end of the cluster). Microscopy images show the green fluorescent channel (GFP) alone and overlaid onto the phase contrast images (merged). Different GFP exposure times (100 ms or 500 ms) are shown to also detect weaker signals. Scale bars: 2 µm.

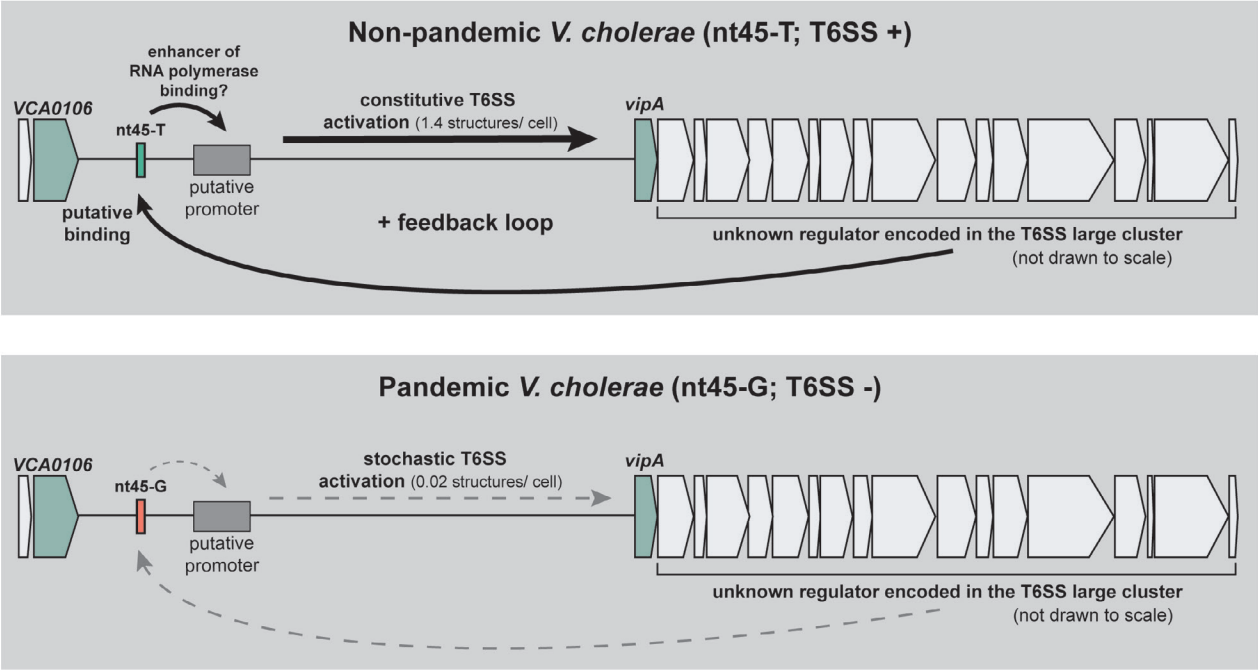
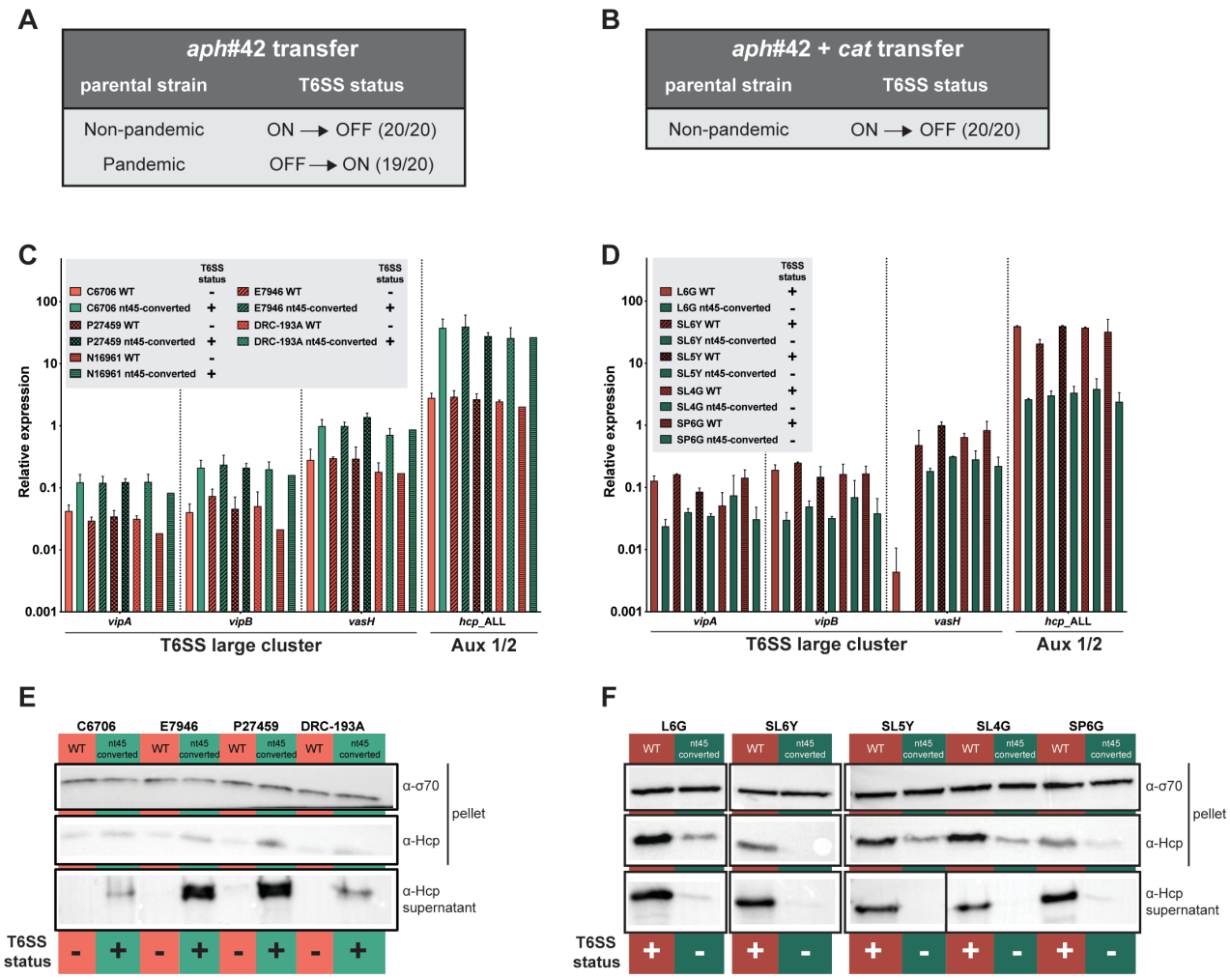


Figure 5.6 Positive feedback loop model of nt45-driven T6SS regulation.

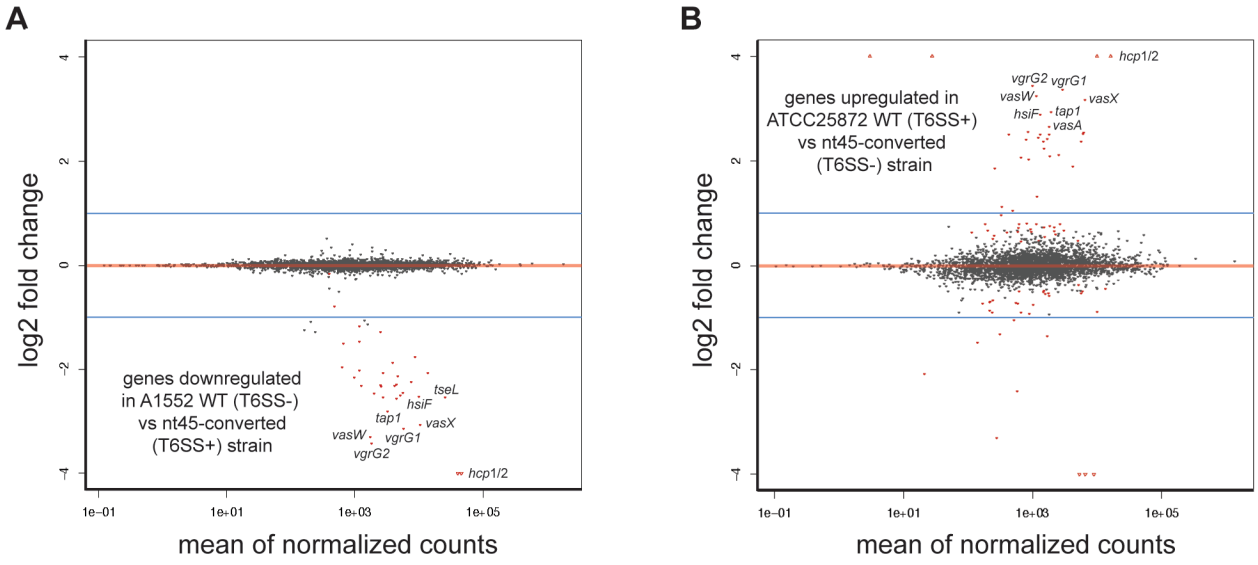
Upper box: nt45 being a 'T' (found in non-pandemic strains; T6SS+) allows transcription of the large T6SS cluster, which putatively contains an unknown regulator that initiates a positive feedback loop. Lower box: with 'G' at the nt45 position (found in pandemic strains; T6SS-), the low transcript levels would rarely result in the production of the regulator, resulting in unspecific T6SS activation in only one in fifty cells while the rest of the population remains T6SS-silent.

5.6 SUPPLEMENTARY FIGURES



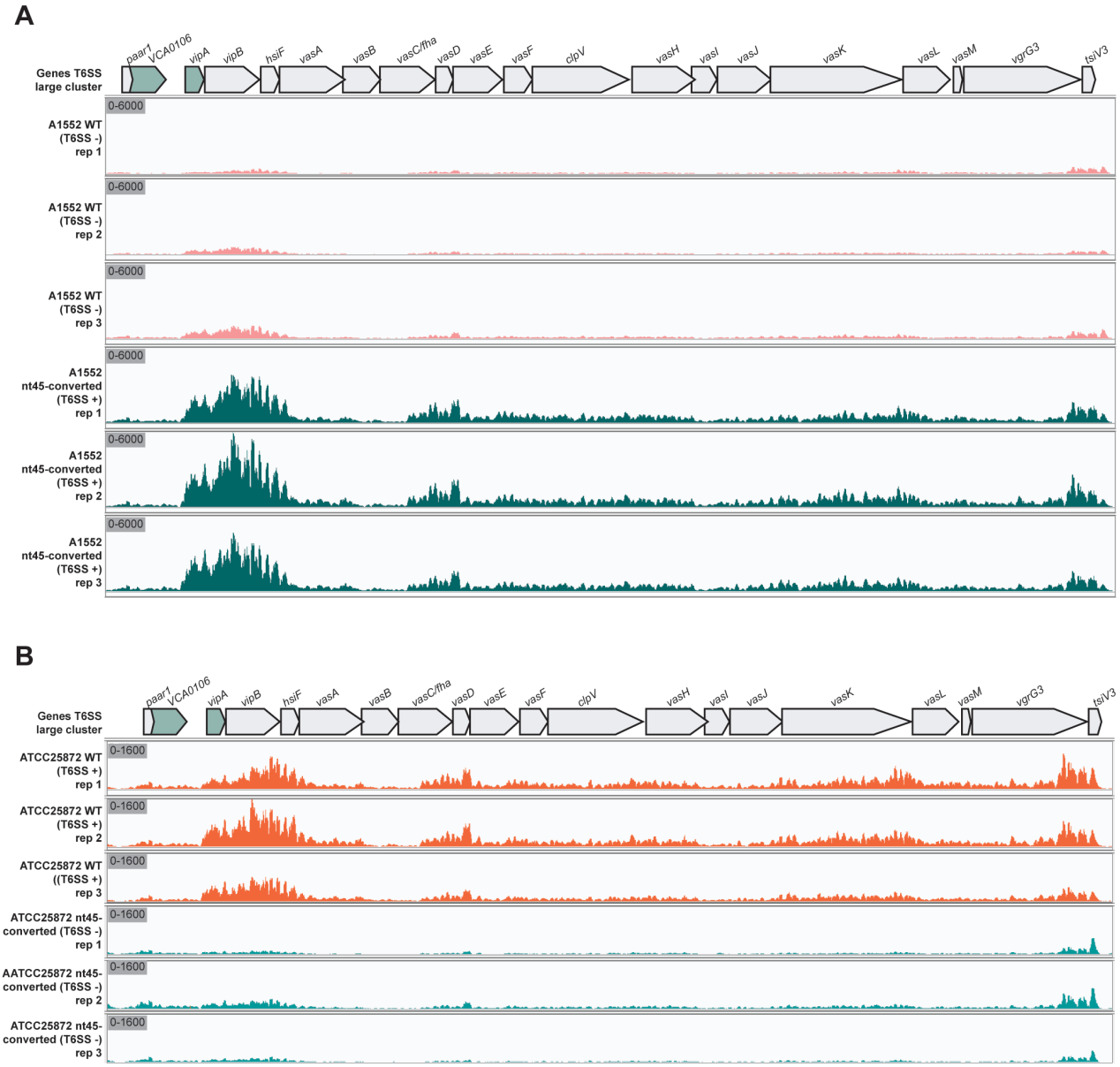
Supplementary figure 5.1 Hybrid clones with transfer of *aph*#42 and *cat* markers and nt45-conversion response in T6SS RNA and protein levels in diverse pandemic and environmental strains.

(A) T6SS activity scores of hybrid clones when the *aph*#42 resistance marker was transferred from a pandemic to a non-pandemic *V. cholerae* strain and vice-versa. **(B)** T6SS activity scores of hybrid clones when the *aph*#42 and *cat* resistance markers were co-transferred (therefore likely including the T6SS large cluster) from a pandemic to a non-pandemic *V. cholerae* strain. **(C-D)** qRT-PCR expression scoring of T6SS genes after growth for 3 h in the WT and nt45-converted variants from pandemic **(C)** and environmental **(D)** *V. cholerae* strains. Reverse primer for *vasH* amplification doesn't bind properly in strain L6G's gene. Data represent the average of two independent biological replicates (±SD). **(E-F)** Hcp protein production and secretion levels in response to nt45-conversion in pandemic (**E**) and environmental (**F**) *V. cholerae* strains after 3 h of growth. Intracellular Hcp (pellet) and secreted (supernatant) Hcp were assessed by immunoblotting using Hcp-directed antibodies. Detection of σ70 on the cell pellets served as a loading control.



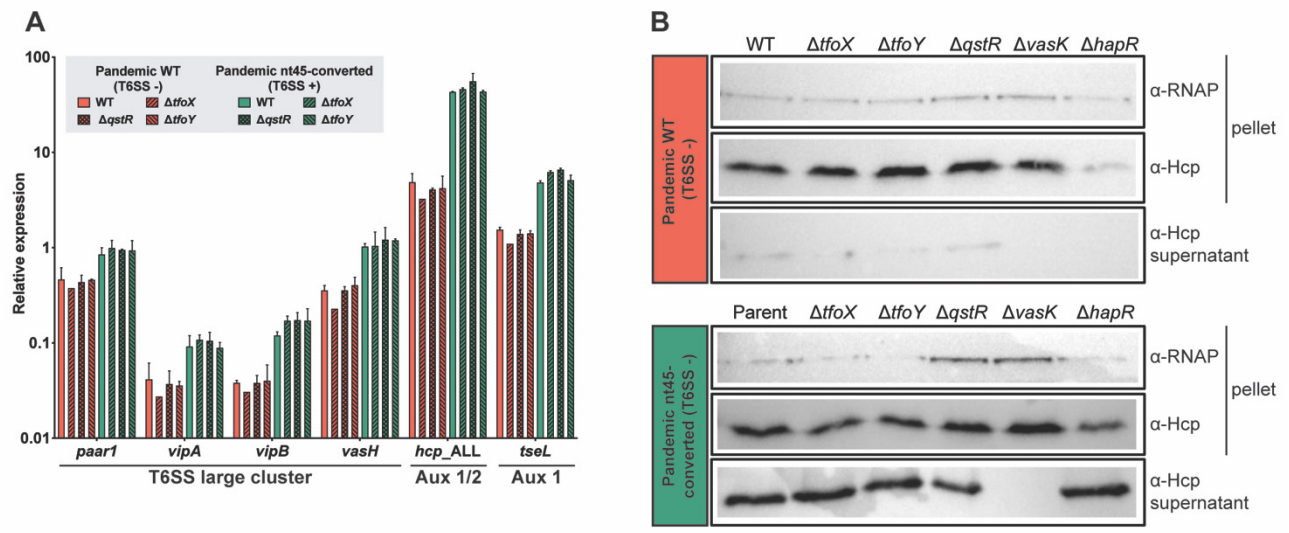
Supplementary figure 5.2 Transcriptional changes derived from nt45 conversion in pandemic and non-pandemic *V. cholerae*.

Log2-fold changes over mean of normalized counts (derived from RNAseq data), when comparing WT and nt45-converted variants of pandemic **(A)** and non-pandemic **(B)** *V. cholerae* strains. Some T6SS genes with the highest log2-fold changes were labelled on the graphs.



Supplementary figure 5.3 Transcript level changes of T6SS large cluster gene in response to nt45 conversion.

RNA-sequencing data showing the coverage of cDNA reads over the large T6SS gene cluster (shown at the top of each panel; *paar1* to *tsiV3*). Data from three independent experiments of the corresponding strains are shown (pandemic T6SS- strain A1552 in **A**; non-pandemic T6SS+ strain ATCC25872 in **B**); the lower three rows show the mapped reads derived from nt45-converted variants of each strain. Y-axis scale ranges from 0-6000 in **A** and from 0-1600 in **B**, as indicated. See Supplementary table 5.2 and Supplementary table 5.3 for details.



Supplementary figure 5.4 Nt45-mediated T6SS regulation is independent of TfoX, TfoY and QstR.

WT and nt45-converted strains carrying or lacking the known T6SS regulators TfoX, TfoY and QstR were checked by qRT-PCR **(A)** and Western blot analysis **(B)** for their T6SS status after growth for 3 h. **(A)** Relative expression of T6SS genes was determined by qRT-PCR. Data represent the average of two independent biological replicates (±SD), except for Δ*tfoX* mutant of pandemic WT strain, which was tested once. **(B)** Intracellular (pellet) and secreted (supernatant) Hcp levels were assessed by immunoblotting using Hcp-directed antibodies. Detection of RNAP served as a loading control of the pellet fraction.

5.7 SUPPLEMENTARY TABLES

Supplementary table 5.1 Bacterial strains and plasmids used in this study.

Strain names	Genotype / description*	Internal strain number	Reference
***V. cholerae* strains**			
A1552	Wild-type, O1 El Tor Inaba, isolated in 1991 in Peru. Rif^R.	MB_1	(66)
A1552-*aph*#1	A1552 (O1) with *aph* cassette in position #1 on chr 1. Rif^R, Kan^R.	MB_6911	This study
A1552-*aph*#2	A1552 (O1) with *aph* cassette in position #2 on chr 1. Rif^R, Kan^R.	MB_6912	This study
A1552-*aph*#3	A1552 (O1) with *aph* cassette in position #3 on chr 1. Rif^R, Kan^R.	MB_6913	This study
A1552-*aph*#4	A1552 (O1) with *aph* cassette in position #4 on chr 1. Rif^R, Kan^R.	MB_6914	This study
A1552-*aph*#5	A1552 (O1) with *aph* cassette in position #5 on chr 1. Rif^R, Kan^R.	MB_6915	This study
A1552-*aph*#6	A1552 (O1) with *aph* cassette in position #6 on chr 1. Rif^R, Kan^R.	MB_6916	This study
A1552-*aph*#7	A1552 (O1) with *aph* cassette in position #7 on chr 1. Rif^R, Kan^R.	MB_6917	This study
A1552-*aph*#8	A1552 (O1) with *aph* cassette in position #8 on chr 1. Rif^R, Kan^R.	MB_6918	This study
A1552-*aph*#9	A1552 (O1) with *aph* cassette in position #9 on chr 1. Rif^R, Kan^R.	MB_6919	This study
A1552-*aph*#10	A1552 (O1) with *aph* cassette in position #10 on chr 1. Rif^R, Kan^R.	MB_6920	This study
A1552-*aph*#11	A1552 (O1) with *aph* cassette in position #11 on chr 1. Rif^R, Kan^R.	MB_6921	This study
A1552-*aph*#12	A1552 (O1) with *aph* cassette in position #12 on chr 1. Rif^R, Kan^R.	MB_6922	This study
A1552-*aph*#13	A1552 (O1) with *aph* cassette in position #13 on chr 1. Rif^R, Kan^R.	MB_6923	This study
A1552-*aph*#14	A1552 (O1) with *aph* cassette in position #14 on chr 1. Rif^R, Kan^R.	MB_6924	This study
A1552-*aph*#15	A1552 (O1) with *aph* cassette in position #15 on chr 1. Rif^R, Kan^R.	MB_6925	This study
A1552-*aph*#16	A1552 (O1) with *aph* cassette in position #16 on chr 1. Rif^R, Kan^R.	MB_6926	This study
A1552-*aph*#17	A1552 (O1) with *aph* cassette in position #17 on chr 1. Rif^R, Kan^R.	MB_6927	This study
A1552-*aph*#18	A1552 (O1) with *aph* cassette in position #18 on chr 1. Rif^R, Kan^R.	MB_6928	This study
A1552-*aph*#19	A1552 (O1) with *aph* cassette in position #19 on chr 1. Rif^R, Kan^R.	MB_6929	This study
A1552-*aph*#20	A1552 (O1) with *aph* cassette in position #20 on chr 1. Rif^R, Kan^R.	MB_6930	This study
A1552-*aph*#21	A1552 (O1) with *aph* cassette in position #21 on chr 1. Rif^R, Kan^R.	MB_6931	This study
A1552-*aph*#22	A1552 (O1) with *aph* cassette in position #22 on chr 1. Rif^R, Kan^R.	MB_6932	This study
A1552-*aph*#23	A1552 (O1) with *aph* cassette in position #23 on chr 1. Rif^R, Kan^R.	MB_6933	This study

A1552-*aph*#24	A1552 (O1) with *aph* cassette in position #24 on chr 1. Rif^R, Kan^R.	MB_6934	This study
A1552-*aph*#25	A1552 (O1) with *aph* cassette in position #25 on chr 1. Rif^R, Kan^R.	MB_6935	This study
A1552-*aph*#26	A1552 (O1) with *aph* cassette in position #26 on chr 1. Rif^R, Kan^R.	MB_6936	This study
A1552-*aph*#27	A1552 (O1) with *aph* cassette in position #27 on chr 1. Rif^R, Kan^R.	MB_6937	This study
A1552-*aph*#28	A1552 (O1) with *aph* cassette in position #28 on chr 1. Rif^R, Kan^R.	MB_6938	This study
A1552-*aph*#29	A1552 (O1) with *aph* cassette in position #29 on chr 1. Rif^R, Kan^R.	MB_6939	This study
A1552-*aph*#30	A1552 (O1) with *aph* cassette in position #30 on chr 1. Rif^R, Kan^R.	MB_6940	This study
A1552-*aph*#31	A1552 (O1) with *aph* cassette in position #31 on chr 2. Rif^R, Kan^R.	MB_6941	This study
A1552-*aph*#32	A1552 (O1) with *aph* cassette in position #32 (15kb upstream of T6SS large cluster) on chr 2. Rif^R, Kan^R.	MB_6942	This study
A1552-*aph*#33	A1552 (O1) with *aph* cassette in position #33 on chr 2. Rif^R, Kan^R.	MB_6943	This study
A1552-*aph*#34	A1552 (O1) with *aph* cassette in position #34 on chr 2. Rif^R, Kan^R.	MB_6944	This study
A1552-*aph*#35	A1552 (O1) with *aph* cassette in position #35 on chr 2. Rif^R, Kan^R.	MB_6945	This study
A1552-*aph*#36	A1552 (O1) with *aph* cassette in position #36 on chr 2. Rif^R, Kan^R.	MB_6946	This study
A1552-*aph*#37	A1552 (O1) with *aph* cassette in position #37 on chr 2. Rif^R, Kan^R.	MB_6947	This study
A1552-*aph*#38	A1552 (O1) with *aph* cassette in position #38 on chr 2. Rif^R, Kan^R.	MB_6948	This study
A1552-*aph*#39	A1552 (O1) with *aph* cassette in position #39 on chr 2. Rif^R, Kan^R.	MB_6949	This study
A1552-*aph*#40	A1552 (O1) with *aph* cassette in position #40 on chr 2. Rif^R, Kan^R.	MB_6950	This study
A1552-*aph*#42	A1552 (O1) with *aph* cassette in position #42 on chr 2 (upstream of *paar1* – *VCA0105*). Rif^R, Kan^R.	MB_7922	This study
A1552-*aph*#32(ATCC25872) hybrid clones 1-20	A1552 (O1) transformed with gDNA of strain ATCC25872-*aph*#32 (#6982). 20 hybrid clones selected; Rif^R, Kan^R.	ND_178 to 197	This study
A1552-T6SS[nt45-C]	A1552 with nt45 in intergenic region between *VCA0106* and *VCA0107* (*vipA*) converted from G to C by site-directed point mutation. Rif^R.	ND_333	This study
A1552-T6SS[nt45-A]	A1552 with nt45 in intergenic region between *VCA0106* and *VCA0107* (*vipA*) converted from G to A by site-directed point mutation. Rif^R.	ND_334	This study
A1552-T6SS[Δ10bp]	A1552 with deletion of 10bp around nt45 in intergenic region between *VCA0106* and *VCA0107* (*vipA*). Rif^R.	ND_335	This study
A1552Δ*tfoX*	A1552 with *tfoX* (*VC1153*) deleted. Rif^R.	MB_45	(27)
A1552Δ*qstR*	A1552 with *qstR* (*VC0396*) deleted. Rif^R.	MB_600	(28)
A1552Δ*tfoY*	A1552 with *tfoY* (*VC1722*) deleted. Rif^R.	MB_828	(17)
A1552Δ*vipB*	A1552 with *vipB* (*VCA0108*) deleted. Rif^R.	MB_598	(30)
A1552Δ*vasK*	A1552 with *vasK* (*VCA0120*) deleted. Rif^R.	MB_585	(30)
A1552Δ*hapR*	A1552 with *hapR* (*VC0583*) deleted. Rif^R.	MB_3	(27)

A1552-Δ300bp-VCA0106-07	A1552 with deletion of 300bp in intergenic region between *VCA0106* and *VCA0107* (*vipA*). Rif^R.	ND_567	This study
A1552-Δ350bp-VCA0106-07	A1552 with deletion of 350bp in intergenic region between *VCA0106* and *VCA0107* (*vipA*). Rif^R.	ND_569	This study
A1552-Δfull-VCA0106-07	A1552 with deletion of intergenic region between *VCA0106* and *VCA0107* (*vipA*). Rif^R.	ND_370	This study
A1552-*vipA-sfGFP*v2	A1552 carrying *vipA-sfGFP* translational fusion (v2: without ATG at start of *sfGFP*); TransFLP method. Rif^R.	MB_3909	This study
A1552-Tn*tfoX*-strep	A1552 containing mini-Tn7-*araC*-P_{BAD}-*tfoX-strep*; Rif^R,Gent^R.	MB_3420	(17)
A1552-Tn*tfoX*-strep-change-10box-VCA0106-07	A1552-Tn*tfoX*-strep with mutation (AA to GC) in putative -10box found in intergenic region between *VCA0106* and *VCA0107* (*vipA*). Rif^R, Gent^R.	ND_768	This study
A1552-Tn*qstR*	A1552 containing mini-Tn7-*araC*-P_{BAD}-*qstR*; Rif^R, Gent^R.	MB_5501	(29)
A1552-Tn*qstR*-change-10box-VCA0106-07	A1552-Tn*qstR* with mutation (AA to GC) in putative -10box found in intergenic region between *VCA0106* and *VCA0107* (*vipA*). Rif^R, Gent^R.	ND_770	This study
A1552-Tn*tfoY*-strep	A1552 containing mini-Tn7-*araC*-P_{BAD}-*tfoY-strep*; Rif^R, Gent^R.	MB_2978	(17)
A1552-Tn*tfoY*-strep - change-10box-VCA0106-07	A1552-Tn*tfoY*-strep with mutation (AA to GC) in putative -10box found in intergenic region between *VCA0106* and *VCA0107* (*vipA*). Rif^R, Gent^R.	ND_769	This study
A1552-T6SS[nt45-T]	A1552 with nt45 in intergenic region between *VCA0106* and *VCA0107* (*vipA*) converted from G to T by site-directed point mutation. Rif^R.	MB_9063	This study
A1552-T6SS [nt45-T] Δ*tfoX*	A1552-T6SS[nt45-T] with *tfoX* (*VC1153*) deleted. Rif^R.	ND_320	This study
A1552-T6SS[nt45-T] Δ*qstR*	A1552-T6SS[nt45-T] with *qstR* (*VC0396*) deleted. Rif^R.	ND_322	This study
A1552-T6SS[nt45-T] Δ*tfoY*	A1552-T6SS[nt45-T] with *tfoY* (*VC1722*) deleted. Rif^R.	ND_321	This study
A1552-T6SS[nt45-T] Δ*vipB*	A1552-T6SS[nt45-T] with *vipB* (*VCA0108*) deleted. Rif^R.	ND_339	This study
A1552-T6SS[nt45-T] Δ*vasK*	A1552-T6SS[nt45-T] with *vasK* (*VCA0120*) deleted. Rif^R.	ND_323	This study
A1552-T6SS[nt45-T] Δ*hapR*	A1552-T6SS[nt45-T] with *hapR* (*VC0583*) deleted. Rif^R.	ND_324	This study
A1552-T6SS[nt45-T] Δ*hapA*	A1552-T6SS[nt45-T] with *hapA*(*VCA0865*) deleted; TransFLP method. Rif^R.	ND_704	This study
A1552-T6SS[nt45-T] Δ*vasK*(full)	A1552-T6SS[nt45-T] with *vasK* (*VCA0120*) deleted (entire ORF). Rif^R.	ND_424	This study
A1552-T6SS[nt45-T] Δ*hapR*Δ*vasK*(full)	A1552-T6SS[nt45-T] with *hapR* (*VC0583*) and *vasK* (*VCA0120;* entire ORF deleted) deleted. Rif^R.	ND_639	This study
A1552-T6SS[nt45-T]-Δ300bp-VCA0106-07	A1552-T6SS[nt45-T] with deletion of 300bp in intergenic region between *VCA0106* and *VCA0107* (*vipA*). Rif^R.	ND_568	This study
A1552-T6SS[nt45-T]-Δ350bp-VCA0106-07	A1552-T6SS[nt45-T] with deletion of 350bp in intergenic region between *VCA0106* and *VCA0107* (*vipA*). Rif^R.	ND_570	This study
A1552-T6SS[nt45P-T]-Δ60bp-VCA0106-07	A1552-T6SS[nt45-T] with deletion of 60bp in intergenic region between *VCA0106* and *VCA0107* (*vipA*). Rif^R.	ND_755	This study
A1552-T6SS[nt45-T]-change-10box-VCA0106-07	A1552-T6SS[nt45-T] with mutation (AA to GC) in putative -10box found in intergenic region between *VCA0106* and *VCA0107* (*vipA*). Rif^R.	ND_757	This study

A1552-T6SS[nt45-T]-*vipA-sfGFP*v2	A1552-T6SS[nt45-T] carrying *vipA-sfGFP* translational fusion (v2: without ATG at start of *sfGFP*), TransFLP method. Rif^R.	ND_273	This study
A1552-T6SS[nt45-T]-*vipA-sfGFP*v2 Δ*hapR*	A1552-T6SS[nt45-T]-*vipA-sfGFP*v2 with *hapR* (*VC0583*) deleted. Rif^R.	ND_641	This study
A1552-T6SS[nt45-T]-*vipA-sfGFP*v2 Δ*vasK*(full)	A1552-T6SS[nt45-T]-*vipA-sfGFP*v2 with *vasK* (*VCA0120*) deleted (entire ORF). Rif^R.	ND_644	This study
A1552Δddm	A1552 with deletion of DNA Defense Module (DDM): ΔVC1770-1772::FRT and ΔVC0490-0493::FRT; TransFLP method. Rif^R.	MB_8560	Jaskólska, Adams *et al.*, in revision
A1552Δddm: pBAD-T6SS[nt45-G]-*vipA-sfGFP*v2	A1552Δddm carrying plasmid pBAD-T6SS[nt45-G]-*vipA-sfGFP*v2. Rif^R, Kan^R.	ND_789	This study
A1552Δddm: pBAD-T6SS[nt45-T]-*vipA-sfGFP*v2	A1552Δddm carrying plasmid pBAD-T6SS[nt45-T]-*vipA-sfGFP*v2. Rif^R, Kan^R.	ND_792	This study
A1552Δddm-*vipA-sfGFP*v2	A1552Δddm carrying *vipA-sfGFP* translational fusion (v2: without ATG at start of *sfGFP*); TransFLP method. Rif^R.	ND_785	This study
A1552Δddm-*vipA-sfGFP*v2: pBAD- T6SS[nt45-G]-*vipA*	A1552Δddm-v*ipA-sfGFP*v2 carrying plasmid pBAD-T6SS[nt45-G]-*vipA*. Rif^R, Kan^R.	ND_793	This study
A1552Δddm-*vipA-sfGFP*v2: pBAD- T6SS[nt45-T]-*vipA*	A1552Δddm-*vipA-sfGFP*v2 carrying plasmid pBAD-T6SS[nt45-T]-*vipA*. Rif^R, Kan^R.	ND_794	This study
A1552Δddm-T6SS [nt45-T]	A1552Δddm with nt45 in intergenic region between *VCA0106* and *VCA0107* (*vipA*) converted from G to T by site-directed point mutation. Rif^R.	ND_786	This study
A1552Δddm-T6SS [nt45-T]: pBAD-T6SS[nt45-G]-*vipA-sfGFP*v2	A1552Δddm-T6SS[nt45-T] carrying plasmid pBAD-T6SS[nt45-G]-*vipA-sfGFP*v2. Rif^R, Kan^R.	ND_795	This study
A1552Δddm-T6SS [nt45-T]: pBAD-T6SS[nt45-T]-*vipA-sfGFP*v2	A1552Δddm-T6SS[nt45-T] carrying plasmid pBAD-T6SS[nt45-T]-*vipA-sfGFP*v2. Rif^R, Kan^R.	ND_796	This study
A1552Δddm-T6SS [nt45-T]Δ*vasH*	A1552Δddm-T6SS[nt45-T] deleted for *vasH* (*VCA0117*); TransFLP method. Rif^R.	ND_810	This study
A1552Δddm-T6SS [nt45-T]Δ*vasH*: pBAD-T6SS[nt45-G]-*vipA-sfGFP*v2	A1552Δddm-T6SS[nt45-T]Δ*vasH* carrying plasmid pBAD-T6SS[nt45-G]-*vipA-sfGFP*v2. Rif^R, Kan^R.	ND_814	This study
A1552Δddm-T6SS [nt45-T]Δ*VCA0122*	A1552Δddm-T6SS[nt45-T] deleted for *VCA0122*; TransFLP method. Rif^R.	ND_813	This study
A1552Δddm-T6SS [nt45-T]Δ*VCA0122*: pBAD-T6SS[nt45-G]-v*ipA-sfGFP*v2	A1552Δddm-T6SS[nt45-T]Δ*VCA0122* carrying plasmid pBAD-T6SS[nt45-G]-*vipA-sfGFP*v2. Rif^R, Kan^R.	ND_815	This study
A1552Δddm-T6SS [nt45-T]Δ*VCA0107-cat*(FRT) Δ*VCA0124-aph*(FRT)	A1552Δddm-T6SS[nt45-T] with *cat* cassette (with FRT sites) inserted in *VCA0107* (*vipA*) and an *aph* cassette (with FRT sites) inserted in *VCA0124* (*tsiV3*). Rif^R, Kan^R, Cm^R.	ND_833	This study
A1552Δddm-T6SS [nt45-T]Δ*VCA0107-0124* (T6SS large cluster)	A1552Δddm-T6SS[nt45-T] deleted for T6SS large cluster (*VCA0107-VCA0124*); TransFLP method. Rif^R.	ND_856	This study
A1552Δddm-T6SS [nt45-T]Δ*VCA0107-0124*: pBAD-T6SS[nt45-G]-*vipA-sfGFP*v2	A1552Δddm-T6SS[nt45-T] deleted for T6SS large cluster (*VCA0107-VCA0124*) carrying plasmid pBAD-T6SS[nt45-G]-*vipA-sfGFP*v2. Rif^R, Kan^R.	ND_857	This study
ATCC25872	Wild-type, non-O1 strain (O37); isolated in 1965 in Czechoslovakia; intermediate resistant to Strep.	MB_276	(40)
ATCC25872-*aph*#32	ATCC25872 (O37) with *aph* cassette in position #32 on chr 2 (15kb upstream of T6SS large cluster); intermediate resistant to Strep, Kan^R.	MB_6982	This study

ATCC25872-*aph*#1(A1552) hybrid clones 1-20	ATCC25872 (O37) transformed with gDNA of strain A1552-*aph*#1 (#6911). 20 hybrid clones selected; intermediate resistant to Strep, Kan^R.	MB_7001 to 7020	This study
ATCC25872-*aph*#2(A1552) hybrid clones 1-20	ATCC25872 (O37) transformed with gDNA of strain A1552-*aph*#2 (#6912). 20 hybrid clones selected; intermediate resistant to Strep, Kan^R.	MB_7021 to 7040	This study
ATCC25872-*aph*#3(A1552) hybrid clones 1-20	ATCC25872 (O37) transformed with gDNA of strain A1552-*aph*#3 (#6913). 20 hybrid clones selected; intermediate resistant to Strep, Kan^R.	MB_7041 to 7060	This study
ATCC25872-*aph*#4(A1552) hybrid clones 1-20	ATCC25872 (O37) transformed with gDNA of strain A1552-*aph*#4 (#6914). 20 hybrid clones selected; intermediate resistant to Strep, Kan^R.	MB_7061 to 7080	This study
ATCC25872-*aph*#5(A1552) hybrid clones 1-20	ATCC25872 (O37) transformed with gDNA of strain A1552-*aph*#5 (#6915). 20 hybrid clones selected; intermediate resistant to Strep, Kan^R.	MB_7081 to 7100	This study
ATCC25872-*aph*#6(A1552) hybrid clones 1-20	ATCC25872 (O37) transformed with gDNA of strain A1552-*aph*#6 (#6916). 20 hybrid clones selected; intermediate resistant to Strep, Kan^R.	MB_7101 to 7120	This study
ATCC25872-*aph*#7(A1552) hybrid clones 1-20	ATCC25872 (O37) transformed with gDNA of strain A1552-*aph*#7 (#6917). 20 hybrid clones selected; intermediate resistant to Strep, Kan^R.	MB_7121 to 7140	This study
ATCC25872-*aph*#8(A1552) hybrid clones 1-20	ATCC25872 (O37) transformed with gDNA of strain A1552-*aph*#8 (#6918). 20 hybrid clones selected; intermediate resistant to Strep, Kan^R.	MB_7141 to 7160	This study
ATCC25872-*aph*#9(A1552) hybrid clones 1-20	ATCC25872 (O37) transformed with gDNA of strain A1552-*aph*#9 (#6919). 20 hybrid clones selected; intermediate resistant to Strep, Kan^R.	MB_7161 to 7180	This study
ATCC25872-*aph*#10(A1552) hybrid clones 1-20	ATCC25872 (O37) transformed with gDNA of strain A1552-*aph*#10 (#6920). 20 hybrid clones selected; intermediate resistant to Strep, Kan^R.	MB_7181 to 7200	This study
ATCC25872-*aph*#11(A1552) hybrid clones 1-20	ATCC25872 (O37) transformed with gDNA of strain A1552-*aph*#11 (#6921). 20 hybrid clones selected; intermediate resistant to Strep, Kan^R.	MB_7201 to 7220	This study
ATCC25872-*aph*#12(A1552) hybrid clones 1-20	ATCC25872 (O37) transformed with gDNA of strain A1552-*aph*#12 (#6922). 20 hybrid clones selected; intermediate resistant to Strep, Kan^R.	MB_7221 to 7240	This study
ATCC25872-*aph*#13(A1552) hybrid clones 1-20	ATCC25872 (O37) transformed with gDNA of strain A1552-*aph*#13 (#6923). 20 hybrid clones selected; intermediate resistant to Strep, Kan^R.	MB_7241 to 7260	This study
ATCC25872-*aph*#14(A1552) hybrid clones 1-20	ATCC25872 (O37) transformed with gDNA of strain A1552-*aph*#14 (#6924). 20 hybrid clones selected; intermediate resistant to Strep, Kan^R.	MB_7261 to 7280	This study
ATCC25872-*aph*#15(A1552) hybrid clones 1-20	ATCC25872 (O37) transformed with gDNA of strain A1552-*aph*#15 (#6925). 20 hybrid clones selected; intermediate resistant to Strep, Kan^R.	MB_7281 to 7300	This study
ATCC25872-*aph*#16(A1552) hybrid clones 1-20	ATCC25872 (O37) transformed with gDNA of strain A1552-*aph*#16 (#6926). 20 hybrid clones selected; intermediate resistant to Strep, Kan^R.	MB_7301 to 7320	This study
ATCC25872-*aph*#17(A1552) hybrid clones 1-20	ATCC25872 (O37) transformed with gDNA of strain A1552-*aph*#17 (#6927). 20 hybrid clones selected; intermediate resistant to Strep, Kan^R.	MB_7321 to 7340	This study
ATCC25872-*aph*#18(A1552) hybrid clones 1-20	ATCC25872 (O37) transformed with gDNA of strain A1552-*aph*#18 (#6928). 20 hybrid clones selected; intermediate resistant to Strep, Kan^R.	MB_7341 to 7360	This study
ATCC25872-*aph*#19(A1552) hybrid clones 1-20	ATCC25872 (O37) transformed with gDNA of strain A1552-*aph*#19 (#6929). 20 hybrid clones selected; intermediate resistant to Strep, Kan^R.	MB_7361 to 7380	This study

ATCC25872-*aph*#20(A1552) hybrid clones 1-20	ATCC25872 (O37) transformed with gDNA of strain A1552-*aph*#20 (#6930). 20 hybrid clones selected; intermediate resistant to Strep, Kan^R.	MB_7381 to 7400	This study
ATCC25872-*aph*#21(A1552) hybrid clones 1-20	ATCC25872 (O37) transformed with gDNA of strain A1552-*aph*#21 (#6931). 20 hybrid clones selected; intermediate resistant to Strep, Kan^R.	MB_7401 to 7420	This study
ATCC25872-*aph*#22(A1552) hybrid clones 1-20	ATCC25872 (O37) transformed with gDNA of strain A1552-*aph*#22 (#6932). 20 hybrid clones selected; intermediate resistant to Strep, Kan^R.	MB_7421 to 7440	This study
ATCC25872-*aph*#23(A1552) hybrid clones 1-20	ATCC25872 (O37) transformed with gDNA of strain A1552-*aph*#23 (#6933). 20 hybrid clones selected; intermediate resistant to Strep, Kan^R.	MB_7441 to 7460	This study
ATCC25872-*aph*#24(A1552) hybrid clones 1-20	ATCC25872 (O37) transformed with gDNA of strain A1552-*aph*#24 (#6934). 20 hybrid clones selected; intermediate resistant to Strep, Kan^R.	MB_7461 to 7480	This study
ATCC25872-*aph*#25(A1552) hybrid clones 1-20	ATCC25872 (O37) transformed with gDNA of strain A1552-*aph*#25 (#6935). 20 hybrid clones selected; intermediate resistant to Strep, Kan^R.	MB_7481 to 7500	This study
ATCC25872-*aph*#26(A1552) hybrid clones 1-20	ATCC25872 (O37) transformed with gDNA of strain A1552-*aph*#26 (#6936). 20 hybrid clones selected; intermediate resistant to Strep, Kan^R.	MB_7501 to 7520	This study
ATCC25872-*aph*#27(A1552) hybrid clones 1-20	ATCC25872 (O37) transformed with gDNA of strain A1552-*aph*#27 (#6937). 20 hybrid clones selected; intermediate resistant to Strep, Kan^R.	MB_7521 to 7540	This study
ATCC25872-*aph*#28(A1552) hybrid clones 1-20	ATCC25872 (O37) transformed with gDNA of strain A1552-*aph*#28 (#6938). 20 hybrid clones selected; intermediate resistant to Strep, Kan^R.	MB_7541 to 7560	This study
ATCC25872-*aph*#29(A1552) hybrid clones 1-20	ATCC25872 (O37) transformed with gDNA of strain A1552-*aph*#29 (#6939). 20 hybrid clones selected; intermediate resistant to Strep, Kan^R.	MB_7561 to 7580	This study
ATCC25872-*aph*#30(A1552) hybrid clones 1-20	ATCC25872 (O37) transformed with gDNA of strain A1552-*aph*#30 (#6940). 20 hybrid clones selected; intermediate resistant to Strep, Kan^R.	MB_7581 to 7600	This study
ATCC25872-*aph*#31(A1552) hybrid clones 1-20	ATCC25872 (O37) transformed with gDNA of strain A1552-*aph*#31 (#6941). 20 hybrid clones selected; intermediate resistant to Strep, Kan^R.	MB_7601 to 7620	This study
ATCC25872-*aph*#32(A1552) hybrid clones 1-20	ATCC25872 (O37) transformed with gDNA of strain A1552-*aph*#32 (#6942). 20 hybrid clones selected; intermediate resistant to Strep, Kan^R.	MB_7621 to 7640	This study
ATCC25872-*aph*#33(A1552) hybrid clones 1-20	ATCC25872 (O37) transformed with gDNA of strain A1552-*aph*#33 (#6943). 20 hybrid clones selected; intermediate resistant to Strep, Kan^R.	MB_7641 to 7660	This study
ATCC25872-*aph*#34(A1552) hybrid clones 1-20	ATCC25872 (O37) transformed with gDNA of strain A1552-*aph*#34 (#6944). 20 hybrid clones selected; intermediate resistant to Strep, Kan^R.	MB_7661 to 7680	This study
ATCC25872-*aph*#35(A1552) hybrid clones 1-20	ATCC25872 (O37) transformed with gDNA of strain A1552-*aph*#35 (#6945). 20 hybrid clones selected; intermediate resistant to Strep, Kan^R.	MB_7681 to 7700	This study
ATCC25872-*aph*#36(A1552) hybrid clones 1-20	ATCC25872 (O37) transformed with gDNA of strain A1552-*aph*#36 (#6946). 20 hybrid clones selected; intermediate resistant to Strep, Kan^R.	MB_7701 to 7720	This study
ATCC25872-*aph*#37(A1552) hybrid clones 1-20	ATCC25872 (O37) transformed with gDNA of strain A1552-*aph*#37 (#6947). 20 hybrid clones selected; intermediate resistant to Strep, Kan^R.	MB_7721 to 7740	This study
ATCC25872-*aph*#38(A1552) hybrid clones 1-20	ATCC25872 (O37) transformed with gDNA of strain A1552-*aph*#38 (#6948). 20 hybrid clones selected; intermediate resistant to Strep, Kan^R.	MB_7741 to 7760	This study

ATCC25872-*aph*#39(A1552) hybrid clones 1-20	ATCC25872 (O37) transformed with gDNA of strain A1552-*aph*#39 (#6949). 20 hybrid clones selected; intermediate resistant to Strep, Kan^R.	MB_7761 to 7780	This study
ATCC25872-*aph*#40(A1552) hybrid clones 1-20	ATCC25872 (O37) transformed with gDNA of strain A1552-*aph*#40 (#6950). 20 hybrid clones selected; intermediate resistant to Strep, Kan^R.	MB_7781 to 7800	This study
ATCC25872-*aph*#42(A1552) hybrid clones 1-20	ATCC25872 (O37) transformed with gDNA of strain A1552-*aph*#42 (#7922). 20 hybrid clones selected; intermediate resistant to Strep, Kan^R.	MB_7951 to 7970	This study
ATCC25872-*cat*#02	ATCC25872 (O37) with *cat* cassette in position #02 on chr 2 (downstream of *tsiV3* – *VCA0124*); intermediate resistant to Strep, Cm^R.	MB_7926	This study
ATCC25872-*cat*#02-*aph*#42(A1552) hybrid clones 1-20	ATCC25872-*cat*#02 transformed with gDNA of strain A1552-*aph*#42 (#7922). 20 hybrid clones selected; intermediate resistant to Strep, Kan^R, Cm^R.	MB_8021 to 8040	This study
ATCC25872-T6SS [nt45-G]-old	ATCC25872 with nt45 in intergenic region between *VCA0106* and *VCA0107* (*vipA*) converted from T to G by site-directed point mutation. Intermediate resistant to Strep.	MB_9166	This study
ATCC25872-*vipA-sfGFP*v2	ATCC25872 carrying *vipA-sfGFP* translational fusion (v2: without ATG at start of *sfGFP*); TransFLP method. Intermediate resistant to Strep.	ND_302	This study
ATCC25872-T6SS [nt45-G]-*vipA-sfGFP*v2	ATCC25872-*vipA-sfGFP*v2 with nt45 in intergenic region between *VCA0106* and *VCA0107* (*vipA*) converted from T to G by site-directed point mutation. Intermediate resistant to Strep.	ND_643	This study
W10G	Wild-type; environmental isolate (non-O1/ non-O139) collected in 2004 in Waddell Creek (CA, USA).	MB_5537	(67)
SA3G	Wild-type; environmental isolate (non-O1/ non-O139) collected in 2004 in Old Salinas River (CA, USA).	MB_957	(67)
SA5Y	Wild-type; environmental isolate (non-O1/ non-O139) collected in 2004 in Old Salinas River (CA, USA).	MB_353	(67)
SL4G	Wild-type; environmental isolate (non-O1/ non-O139) collected in 2004 in San Lorenzo River (CA, USA). Amp^R.	MB_955	(67)
SL4G-T6SS[nt45-G]	SL4G with nt45 in intergenic region between *VCA0106* and *VCA0107* (*vipA*) converted from T to G by site-directed point mutation. Amp^R.	ND_513	This study
SL5Y	Wild-type; environmental isolate (non-O1/ non-O139) collected in 2004 in San Lorenzo River (CA, USA).	MB_954	(67)
SL5Y-T6SS[nt45-G]	SL5Y with nt45 in intergenic region between *VCA0106* and *VCA0107* (*vipA*) converted from T to G by site-directed point mutation.	ND_512	This study
SO5Y	Wild-type; environmental isolate (non-O1/ non-O139) collected in 2004 in Soquel Creek (CA, USA).	MB_960	(67)
L6G	Wild-type; environmental isolate (non-O1/ non-O139) collected in 2004 in Lagunitas Creek (CA, USA). Amp^R.	MB_956	(67)
L6G-T6SS[nt45-G]	L6G with nt45 in intergenic region between *VCA0106* and *VCA0107* (*vipA*) converted from T to G by site-directed point mutation. Amp^R.	ND_509	This study

SL6Y	Wild-type; environmental isolate (non-O1/ non-O139) collected in 2004 in San Lorenzo River (CA, USA).	MB_953	(67)
SL6Y-T6SS[nt45-G]	SL6Y with nt45 in intergenic region between *VCA0106* and *VCA0107* (*vipA*) converted from T to G by site-directed point mutation.	ND_511	This study
SP6G	Wild-type; environmental isolate (non-O1/ non-O139) collected in 2004 in San Pedro Creek (CA, USA).	MB_964	(67)
SP6G-T6SS[nt45-G]	SP6G with nt45 in intergenic region between *VCA0106* and *VCA0107* (*vipA*) converted from T to G by site-directed point mutation.	ND_514	This study
SP7G	Wild-type; environmental isolate (non-O1/ non-O139) collected in 2004 in San Pedro Creek (CA, USA).	MB_952	(67)
W6G	Wild-type; environmental isolate (non-O1/ non-O139) collected in 2004 in Waddell Creek (CA, USA).	MB_354	(67)
W7G	Wild-type; environmental isolate (non-O1/ non-O139) collected in 2004 in Waddell Creek (CA, USA).	MB_962	(67)
E7G	Wild-type; environmental isolate (non-O1/ non-O139) collected in 2004 in Moss Landing Harbor (CA, USA).	MB_963	(67)
SA7G	Wild-type; environmental isolate (non-O1/ non-O139) collected in 2004 in Old Salinas River (CA, USA).	MB_959	(67)
SA10G	Wild-type; environmental isolate (non-O1/ non-O139) collected in 2004 in Old Salinas River (CA, USA).	MB_5539	(67)
C6706 (Strep^{S}) (original)	Wild-type; O1 El Tor Inaba collected in 1991 in Peru. Original isolate before introduction of streptomycin resistance mutation; non-mutated *luxO.* Strep^{S}.	MB_4522	Gift from J. Mekalanos (Harvard); (35)
C6706-T6SS[nt45-T]	C6706 with nt45 in intergenic region between *VCA0106* and *VCA0107* (*vipA*) converted from G to T by site-directed point mutation. Strep^{S}.	ND_538	This study
E7946	Wild-type; O1 El Tor Ogawa isolated in 197 in Bahrain. Strep^{R}.	MB_2600	Lab stock; (68)
E7946-T6SS[nt45-T]	E7946 with nt45 in intergenic region between *VCA0106* and *VCA0107* (*vipA*) converted from G to T by site-directed point mutation. Strep^{R}.	ND_539	This study
P27459	Wild-type; O1 El Tor Inaba isolated in 1976 in Bangladesh. Strep^{R}.	MB_1504	(69)
P27459-T6SS[nt45-T]	P27459 with nt45 in intergenic region between *VCA0106* and *VCA0107* (*vipA*) converted from G to T by site-directed point mutation. Strep^{R}.	ND_540	This study
DRC-193A	Wild-type; O1 isolated in 2011 in the Democratic Republic of Congo. Strep^{R}.	MB_1954	(30)
DRC-193A -T6SS [nt45-T]	DRC-193A with nt45 in intergenic region between *VCA0106* and *VCA0107* (*vipA*) converted from G to T by site-directed point mutation. Strep^{R}.	ND_541	This study
N16961-*hapR*^{Rep}	Wild-type; O1 El Tor Inaba isolated in 1975 in Bangladesh. *hapR* frameshift mutation repaired. Strep^{R}.	MB_5663	(35)

N16961-*hapR*^{Rep}-T6SS [nt45-T]	N16961-*hapR*^{Rep} with nt45 in intergenic region between *VCA0106* and *VCA0107* (*vipA*) converted from G to T by site-directed point mutation. Strep^R.	ND_537	This study
***E. coli* strains**			
TOP10	F- mcrA Δ(mrr-hsdRMS-mcrBC) φ80lacZΔM15 ΔlacX74 nupG recA1 araΔ139 Δ(ara-leu)7697 galE15 galK16 rpsL(StrR) endA1λ-.	MB_741	Invitrogen
TOP10-TnKan	TOP10 containing mini-Tn7-*aph* (Kan^R); Strep^R, Kan^R, Gent^R.	MB_4119	(17)
TOP10-TnGFP	TOP10 containing mini-Tn7-GFP; Strep^R, Cm^R, Gent^R.	MB_4482	This study
MC4100-TnGFP	MC4100 containing mini-Tn7-GFP; Strep^R, Cm^R, Gent^R.	MB_3930	This study
DH5α	F- endA1 glnV44 thi-1 recA1 relA1 gyrA96 deoR nupG φ80lacZΔM15 Δ(lacZYA-argF) U169 hsdR17 (rK$^-$ mK$^+$) phoA, λ-.	MB_736	(70)
S17-1λpir	Tpr Smr recA thi pro hsdR2M1 RP4:2-Tc:Mu:Kmr Tn7 (λpir); Strep^R.	MB_648	(71)
Plasmids			
pBR-FRT-Kan-FRT2	pBR322 derivative containing improved FRT-*aph*-FRT cassette, used as template for TransFLP; Amp^R, Kan^R.	MB_3782	(17)
pBR-FRT-Cat-FRT2	pBR322 derivative containing improved FRT-*cat*-FRT cassette, used as template for TransFLP; Amp^R, Cm^R.	MB_3783	(17)
pBR-flp	pBR322 derivative containing FLP+, λ cI857+, λ pR from pCP20 integrated into the *EcoRV* site of pBR322, used for FLP recombination; Amp^R.	MB_1203	(72)
pGP704-Sac28	Suicide plasmid, *oriR6K sacB* Amp^R.	MB_694	(58)
pGP704-Sac28-VCA0106-VCA0107[nt45-T]	pGP704-Sac28 carrying a genome fragment resulting in a site-directed point mutation in the nt45 (T) located in the intergenic region between *VCA0106* and *VCA0107* (*vipA*). Amp^R.	MB_9062	This study
pGP704-Sac28-VCA0106-VCA0107[nt45-G]	pGP704-Sac28 carrying a genome fragment resulting in a site-directed point mutation in the nt45 (G) located in the intergenic region between *VCA0106* and *VCA0107* (*vipA*). Amp^R.	ND_251	This study
pGP704-Sac28-VCA0106-VCA0107[nt45-C]	pGP704-Sac28 carrying a genome fragment resulting in a site-directed point mutation in the nt45 (C) located in the intergenic region between *VCA0106* and *VCA0107* (*vipA*). Amp^R.	ND_330	This study
pGP704-Sac28-VCA0106-VCA0107[nt45-A]	pGP704-Sac28 carrying a genome fragment resulting in a site-directed point mutation in the nt45 (A) located in the intergenic region between *VCA0106* and *VCA0107* (*vipA*). Amp^R.	ND_331	This study
pGP704-Sac28-VCA0106-VCA0107[Δ10bp]	pGP704-Sac28 carrying a genome fragment resulting in a deletion of 10bp around nt45 located in the intergenic region between *VCA0106* and *VCA0107* (*vipA*). Amp^R.	ND_332	This study
pGP704-Sac28-SL6Y-VCA0106-VCA0107 [nt45-G]	pGP704-Sac28 carrying a genome fragment resulting in a site-directed point mutation in the nt45 (G) located in the intergenic region between *VCA0106* and *VCA0107* (*vipA*) of strain SL6Y. Amp^R.	ND_505	This study

pGP704-Sac28-SP6G-VCA0106-VCA0107 [nt45-G]	pGP704-Sac28 carrying a genome fragment resulting in a site-directed point mutation in the nt45 (G) located in the intergenic region between *VCA0106* and *VCA0107* (*vipA*) of strain SP6G. Amp^R.	ND_506	This study
pGP704-Sac28-Δ*tfoX*	pGP704-Sac28 carrying a gene fragment resulting in a deletion within *tfoX* (*VC1153*). Amp^R.	MB_1013	(17)
pGP704-Sac28-Δ*qstR*	pGP704-Sac28 carrying a gene fragment resulting in a deletion within *qstR* (*VC0396*). Amp^R.	MB_1118	(28)
pGP704-Sac28-Δ*tfoY*	pGP704-Sac28 carrying a gene fragment resulting in a deletion within *tfoY* (*VC1722*). Amp^R.	MB_4116	(17)
pGP704-Sac28-Δ*vipB*	pGP704-Sac28 carrying a gene fragment resulting in a deletion within *vipB* (*VCA0108*). Amp^R.	MB_1123	(30)
pGP704-Sac28-Δ*vasK*	pGP704-Sac28 carrying a gene fragment resulting in a deletion within *vasK* (*VCA0120*). Amp^R.	MB_1124	(30)
pGP704-Sac28-Δ*hapR*	pGP704-Sac28 carrying a gene fragment resulting in a deletion within *hapR* (*VC0583*). Amp^R.	MB_1038	(27)
pGP704-Sac28-Δ*vasK*(full)	pGP704-Sac28 carrying a gene fragment resulting in an almost complete deletion of *vasK* (*VCA0120*). Amp^R.	ND_421	This study
pGP704-Sac28-Δ300bp-VCA0106-07[nt45-G]	pGP704-Sac28 carrying a gene fragment resulting in a ~300bp deletion in the intergenic region between *VCA0106* and *VCA0107* (*vipA*), with a nt45-G. 10bp downstream of *VCA0106* and 10bp upstream of *vipA* were kept. Amp^R.	ND_559	This study
pGP704-Sac28-Δ300bp-VCA0106-07[nt45-T]	pGP704-Sac28 carrying a gene fragment resulting in a ~300bp deletion in the intergenic region between *VCA0106* and *VCA0107* (*vipA*), with a nt45-T. 10bp downstream of *VCA0106* and 10bp upstream of *vipA* were kept. Amp^R.	ND_560	This study
pGP704-Sac28-Δ350bp-VCA0106-07[nt45-G]	pGP704-Sac28 carrying a gene fragment resulting in a ~350bp deletion in the intergenic region between *VCA0106* and *VCA0107* (*vipA*), with a nt45-G. 10bp downstream of *VCA0106* and 10bp upstream of *vipA* were kept. Amp^R.	ND_561	This study
pGP704-Sac28-Δ350bp-VCA0106-07[nt45-T]	pGP704-Sac28 carrying a gene fragment resulting in a ~350bp deletion in the intergenic region between *VCA0106* and *VCA0107* (*vipA*), with a nt45-T. 10bp downstream of *VCA0106* and 10bp upstream of *vipA* were kept. Amp^R.	ND_562	This study
pGP704-Sac28-Δfull-VCA0106-07	pGP704-Sac28 carrying a gene fragment resulting in the deletion of the intergenic region between *VCA0106* and *VCA0107* (*vipA*). 10bp downstream of *VCA0106* and 10bp upstream of *vipA* were kept. Amp^R.	ND_368	This study
pGP704-Sac-Kan	Suicide plasmid, *oriR6K sacB*, Kan^R.	MB_6038	(73)
pGP704-Sac-Kan-L6G-VCA0106-VCA0107 [nt45-G]	pGP704-Sac-Kan carrying a genome fragment resulting in a site-directed point mutation in the nt45 (G) located in the intergenic region between *VCA0106* and *VCA0107* (*vipA*) of strain L6G. Kan^R.	ND_503	This study
pGP704-Sac-Kan-SL5Y-VCA0106-VCA0107 [nt45-G]	pGP704-Sac-Kan carrying a genome fragment resulting in a site-directed point mutation in the nt45 (G) located in the intergenic region between *VCA0106* and *VCA0107* (*vipA*) of strain SL5Y. Kan^R.	ND_506	This study

pGP704-Sac-Kan-SL4G-VCA0106-VCA0107 [nt45-G]	pGP704-Sac-Kan carrying a genome fragment resulting in a site-directed point mutation in the nt45 (G) located in the intergenic region between *VCA0106* and *VCA0107* (*vipA*) of strain SL4G. Kan^R.	ND_507	This study
pGP704-Sac-Kan-Δ60bp-VCA0106-07[nt45-T]	pGP704-Sac-Kan carrying a genome fragment resulting in a ~60bp deletion in the intergenic region between *VCA0106* and *VCA0107* (*vipA*), with a nt45-T. 10bp downstream of *VCA0106* and 10bp upstream of *vipA* were kept. Kan^R.	ND_754	This study
pGP704-Sac-Kan-Δ60bp-VCA0106-07[nt45-G]	pGP704-Sac-Kan carrying a genome fragment resulting in a ~60bp deletion in the intergenic region between *VCA0106* and *VCA0107* (*vipA*), with a nt45-G. 10bp downstream of *VCA0106* and 10bp upstream of *vipA* were kept. Kan^R.	ND_763	This study
pGP704-Sac-Kan- change-10box-VCA0106-07[nt45-T]	pGP704-Sac-Kan carrying a genome fragment resulting in a site-directed point mutation in the putative -10box (AA to GC) located in the intergenic region between *VCA0106* and *VCA0107* (*vipA*)*, with a nt45-T. Kan^R.	ND_756	This study
pGP704-Sac-Kan- change-10box-VCA0106-07[nt45-G]	pGP704-Sac-Kan carrying a genome fragment resulting in a site-directed point mutation in the putative -10box (AA to GC) located in the intergenic region between *VCA0106* and *VCA0107* (*vipA*)*, with a nt45-G. Kan^R.	ND_767	This study
pBAD(Kan)	*bla* replaced by *aph* in pBAD/Myc-HisA; promoter region and MCS maintained; Kan^R.	MB_3363	(50)
pBAD(Kan)-T6SS [nt45-G]-*vipA-sfGFP*v2	pBAD(Kan) carrying region starting 1kb upstream of *VCA0105* (*paar1*) until the end of the translational fusion *vipA-sfGFP*v2, with nt45-G in the intergenic region between *VCA0106* and *VCA0107* (*vipA*). Kan^R.	ND_776	This study
pBAD(Kan)-T6SS [nt45-T]-*vipA-sfGFP*v2	pBAD(Kan) carrying region starting 1kb upstream of *VCA0105* (*paar1*) until the end of the translational fusion *vipA-sfGFP*v2, with nt45-T in the intergenic region between *VCA0106* and *VCA0107* (*vipA*). Kan^R.	ND_777	This study
pBAD(Kan)-T6SS [nt45-G]-*vipA*	pBAD(Kan) carrying region starting 1kb upstream of *VCA0105* (*paar1*) until the end of *VCA0107* (*vipA*), with nt45-G in the intergenic region between *VCA0106* and *VCA0107* (*vipA*). Kan^R.	ND_778	This study
pBAD(Kan)-T6SS [nt45-T]-*vipA*	pBAD(Kan) carrying region starting 1kb upstream of *VCA0105* (*paar1*) until the end of *VCA0107* (*vipA*), with nt45-T in the intergenic region between *VCA0106* and *VCA0107* (*vipA*). Kan^R.	ND_779	This study

*reference locus tags belong to reference strain N16961 according to (74).

Supplementary table 5.2 Relative expression data (normalized counts) for T6SS genes of WT and nt45-converted variants of pandemic strain A1552.

		A1552 WT (T6SS-)				A1552 nt45-converted (T6SS+)				Ratio
	Gene	Repl 1	Repl 2	Repl 3	Average	Repl 1	Repl 2	Repl 3	Average	SNP/ WT
T6SS large cluster	*paar*	373	312	345	343	633	587	657	625	2
	VCA0106	513	625	688	608	1524	1676	2071	1757	3
	vipA	894	2136	2047	1692	7708	11146	10667	9840	6
	vipB	4528	8973	8484	7329	36769	53118	44061	44649	6
	hsiF	1098	1392	1420	1303	7324	8946	7156	7809	6
	vasA	1264	1701	1620	1528	7924	10589	8130	8881	6
	vasB	372	392	487	417	1976	2253	2132	2120	5
	fha	1457	1409	1268	1378	6963	8423	5707	7031	5
	vasD	908	828	795	844	4013	5251	3742	4335	5
	vasE	1539	1467	1512	1506	7689	8454	6329	7490	5
	vasF	936	741	836	838	4349	4681	3675	4235	5
	clpV	2715	2539	2747	2667	12343	14252	11757	12784	5
	vasH	1849	1569	1762	1727	7549	8376	7187	7704	4
	vasI	262	240	275	259	998	1116	975	1029	4
	vasJ	1003	984	1170	1052	3911	5173	4540	4541	4
	vasK	5358	5123	5155	5212	18885	26705	21838	22476	4
	vasL	2165	1330	1313	1603	6299	6759	5504	6187	4
	vgrG3	4508	3634	3470	3870	11642	17215	12326	13727	4
	tsiV3	1974	1070	1074	1372	3724	3988	3313	3675	3
T6SS Aux1	*hcp1*	3889	3752	4344	3995	68835	81595	83018	77816	19
	vgrG1	1037	1096	1297	1144	7125	14324	9930	10460	9
	tap1	808	763	831	801	4518	7661	5163	5781	7
	tseL	3377	2611	2667	2885	14046	22291	15231	17189	6
	tsiV1	718	515	564	599	2692	4325	3247	3421	6
	VC1420	422	293	336	350	1448	1900	1492	1613	5
T6SS Aux2	*hcp2*	4203	4275	5147	4542	75356	92647	97759	88587	20
	vgrG2	226	339	357	307	2695	4129	3361	3395	11
	vasW	299	321	340	320	2710	3929	3029	3223	10
	vasX	2477	2120	2135	2244	16717	23314	17475	19169	9
	tsiV2	1003	649	737	796	4876	4820	4606	4767	6
Aux3	*paar2*	975	458	584	672	1471	1713	1914	1699	3
	tseH	626	316	376	439	1344	2119	2338	1933	4
	tsiH	411	250	319	327	766	1008	1227	1000	3

Supplementary table 5.3 Relative expression data (normalized counts) for T6SS genes of WT and nt45-converted variants of non-pandemic strain ATCC25872.

		ATCC25872 WT (T6SS+)				ATCC25872 nt45-converted (T6SS-)				Ratio
	Gene	Repl 1	Repl 2	Repl 3	Average	Repl 1	Repl 2	Repl 3	Average	WT/ SNP
T6SS large cluster	*paar*	385	359	385	376	291	308	292	297	1
	VCA0106	706	783	847	779	484	574	548	536	1
	vipA	1544	3053	2929	2509	442	545	525	504	5
	vipB	8280	12722	10784	10595	1396	2030	1857	1761	6
	hsiF	2259	2626	2117	2334	247	345	332	308	8
	vasA	2803	3579	2947	3109	367	571	497	478	7
	vasB	734	652	804	730	98	136	141	125	6
	fha	2702	2296	1802	2267	342	446	376	388	6
	vasD	1635	1662	1132	1476	251	259	222	244	6
	vasE	2605	2778	2041	2475	426	527	438	464	5
	vasF	1576	1380	1098	1351	235	280	224	246	5
	clpV	4491	4473	3552	4172	890	1061	867	940	4
	vasH	3205	3152	2592	2983	666	738	650	684	4
	vasI	433	396	394	408	84	120	121	108	4
	vasJ	1449	1456	1340	1415	301	376	341	340	4
	vasK	10261	10150	8146	9519	1560	2016	1793	1789	5
	vasL	3758	2879	2001	2879	524	553	473	516	6
	vgrG3	7884	6578	5368	6610	1783	1917	1459	1720	4
	tsiV3	2256	1586	1222	1688	669	664	561	632	3
T6SS Aux1	*hcp1*	19554	19653	18391	19199	684	510	569	588	33
	vgrG1	5045	5651	5373	5356	487	571	484	514	10
	tap1	3486	3725	3039	3417	424	504	385	438	8
	tseL	11535	10718	8490	10248	1823	1887	1532	1747	6
	tsiV1	2237	2221	1738	2065	434	365	304	368	6
	VC1420	1180	1158	906	1082	281	248	226	252	4
T6SS Aux2	*hcp2*	34594	31399	28601	31531	1071	895	908	958	33
	vgrG2	1692	1789	2042	1841	198	164	141	168	11
	vasW	2132	2153	1952	2079	269	205	175	216	10
	vasX	13234	12176	9993	11801	1498	1267	1103	1289	9
	tsiV2	3451	2862	2849	3054	601	518	468	529	6
Aux3	*paar2*	974	926	867	922	684	645	470	600	2
	tseH	765	907	788	820	544	446	284	424	2
	tsiH	472	569	488	510	385	277	225	296	2

5.8 REFERENCES

1. Boyd,E.F., Carpenter,M.R., Chowdhury,N., Cohen,A.L., Haines-Menges,B.L., Kalburge,S.S., Kingston,J.J., Lubin,J.B., Ongagna-Yhombi,S.Y. and Whitaker,W.B. (2015) Post-genomic analysis of members of the family Vibrionaceae. *Microbiol. Spectr.*, **3**, VE-0009-2014.

2. Dikow,R.B. and Smith,W.L. (2013) Genome-level homology and phylogeny of Vibrionaceae (Gammaproteobacteria: Vibrionales) with three new complete genome sequences. *BMC Microbiol.*, **13**.

3. Granato,E.T., Meiller-Legrand,T.A. and Foster,K.R. (2019) The evolution and ecology of bacterial warfare. *Curr. Biol.*, **29**, R521–R537.

4. Hibbing,M.E., Fuqua,C., Parsek,M.R. and Peterson,S.B. (2010) Bacterial competition: surviving and thriving in the microbial jungle. *Nat. Rev. Microbiol.*, **8**, 15–25.

5. Peterson,S.B., Bertolli,S.K. and Mougous,J.D. (2020) The central role of interbacterial antagonism in bacterial fife. *Curr. Biol.*, **30**, R1203–R1214.

6. Basler,M. (2015) Type VI secretion system: secretion by a contractile nanomachine. *Philos. Trans. R. Soc. B Biol. Sci.*, **370**, 20150021.

7. Cherrak,Y., Flaugnatti,N., Durand,E., Journet,L. and Cascales,E. (2019) Structure and regulation of the type VI secretion system. *Microbiol. Spectr.*, **7**, PSIB-0031-2019.

8. Hachani,A., Wood,T.E. and Filloux,A. (2016) Type VI secretion and anti-host effectors. *Curr. Opin. Microbiol.*, **29**, 81–93.

9. Hood,R.D., Singh,P., Hsu,F.S., Güvener,T., Carl,M.A., Trinidad,R.R.S., Silverman,J.M., Ohlson,B.B., Hicks,K.G., Plemel,R.L., *et al.* (2010) A type VI secretion system of *Pseudomonas aeruginosa* targets a toxin to bacteria. *Cell Host Microbe*, **7**, 25–37.

10. Russell,A.B., Peterson,S.B. and Mougous,J.D. (2014) Type VI secretion system effectors: poisons with a purpose. *Nat. Rev. Microbiol.*, **12**, 137–148.

11. Russell,A.B., Hood,R.D., Bui,N.K., Leroux,M., Vollmer,W. and Mougous,J.D. (2011) Type VI secretion delivers bacteriolytic effectors to target cells. *Nature*, **475**, 343–349.

12. Brooks,T.M., Unterweger,D., Bachmann,V., Kostiuk,B. and Pukatzki,S. (2013) Lytic activity of the *Vibrio cholerae* type VI secretion toxin VgrG-3 is inhibited by the antitoxin TsaB. *J. Biol. Chem.*, **288**, 7618–7625.

13. Dong,T.G., Ho,B.T., Yoder-Himes,D.R. and Mekalanos,J.J. (2013) Identification of T6SS-dependent effector and immunity proteins by Tn-seq in *Vibrio cholerae*. *Proc. Natl. Acad. Sci.*, **110**, 2623–2628.

14. Unterweger,D., Miyata,S.T., Bachmann,V., Brooks,T.M., Mullins,T., Kostiuk,B., Provenzano,D. and Pukatzki,S. (2014) The *Vibrio cholerae* type VI secretion system employs diverse effector modules for intraspecific competition. *Nat. Commun.*, **5**, 3549.

15. Crisan,C. V and Hammer,B.K. (2020) The *Vibrio cholerae* type VI secretion system: toxins, regulators and consequences. *Environ. Microbiol.*, **22**, 4112–4122.

16. Zheng,J., Ho,B. and Mekalanos,J.J. (2011) Genetic analysis of anti-amoebae and anti-bacterial activities of the type VI secretion system in *Vibrio cholerae*. *PLoS One*, **6**, e23876.

17. Metzger,L.C., Stutzmann,S., Scrignari,T., Van der Henst,C., Matthey,N. and Blokesch,M. (2016) Independent regulation of type VI secretion in *Vibrio cholerae* by TfoX and TfoY. *Cell Rep.*, **15**, 951–958.

18. Liang,X., Moore,R., Wilton,M., Wong,M.J.Q., Lam,L. and Dong,T.G. (2015) Identification of divergent type VI secretion effectors using a conserved chaperone domain. *Proc. Natl. Acad. Sci.*, **112**, 9106–9111.

19. Unterweger,D., Kostiuk,B., Otjengerdes,R., Wilton,A., Diaz-Satizabal,L. and Pukatzki,S. (2015) Chimeric adaptor proteins translocate diverse type VI secretion system effectors in *Vibrio cholerae*. *EMBO J.*, **34**, 2198–2210.

20. Russell,A.B., Leroux,M., Hathazi,K., Agnello,D.M., Ishikawa,T., Wiggins,P.A., Wai,S.N. and Mougous,J.D. (2013) Diverse type VI secretion phospholipases are functionally plastic antibacterial effectors. *Nature*, **496**, 508–512.

21. Miyata,S.T., Kitaoka,M., Brooks,T.M., McAuley,S.B. and Pukatzki,S. (2011) *Vibrio cholerae* requires the type VI secretion system virulence factor VasX to kill *Dictyostelium discoideum*. *Infect. Immun.*, **79**, 2941–2949.

22. Miyata,S.T., Unterweger,D., Rudko,S.P. and Pukatzki,S. (2013) Dual expression profile of type VI secretion system immunity genes protects pandemic *Vibrio cholerae*. *PLoS Pathog.*, **9**, e1003752.

23. Santoriello,F.J., Michel,L., Unterweger,D. and Pukatzki,S. (2020) Pandemic *Vibrio cholerae* shuts down site-specific recombination to retain an interbacterial defence mechanism. *Nat. Commun.*, **11**, 6246.

24. Shneider,M.M., Buth,S.A., Ho,B.T., Basler,M., Mekalanos,J.J. and Leiman,P.G. (2013) PAAR-repeat proteins sharpen and diversify the type VI secretion system spike. *Nature*, **500**, 350–353.

25. Hersch,S.J., Watanabe,N., Stietz,M.S., Manera,K., Kamal,F., Burkinshaw,B., Lam,L., Pun,A., Li,M., Savchenko,A., *et al.* (2020) Envelope stress responses defend against type six secretion system attacks independently of immunity proteins. *Nat. Microbiol.*, **5**, 706–714.

26. Altindis,E., Dong,T., Catalano,C. and Mekalanos,J.J. (2015) Secretome analysis of *Vibrio cholerae* type VI secretion system reveals a new effector-immunity pair. *MBio*, **6**, e00075-15.

27. Meibom,K.L., Blokesch,M., Dolganov,N.A., Wu,C.Y. and Schoolnik,G.K. (2005) Chitin induces natural competence in *Vibrio cholerae*. *Science*, **310**, 1824–1827.

28. Lo Scrudato,M. and Blokesch,M. (2013) A transcriptional regulator linking quorum sensing and chitin induction to render *Vibrio cholerae* naturally transformable. *Nucleic Acids Res.*, **41**, 3644–3658.

29. Jaskólska,M., Stutzmann,S., Stoudmann,C. and Blokesch,M. (2018) QstR-dependent regulation of natural competence and type VI secretion in *Vibrio cholerae*. *Nucleic Acids Res.*, **46**, 10619–10634.

30. Borgeaud,S., Metzger,L.C., Scrignari,T. and Blokesch,M. (2015) The type VI secretion system of *Vibrio cholerae* fosters horizontal gene transfer . *Science*, **347**, 63–67.

31. Pollack-Berti,A., Wollenberg,M.S. and Ruby,E.G. (2010) Natural transformation of *Vibrio fischeri* requires tfoX and tfoY. *Environ. Microbiol.*, **12**, 2302–2311.

32. Sudarsan,N., Lee,E.R., Weinberg,Z., Moy,R.H., Kim,J.N., Link,K.H. and Breaker,R.R. (2008) Riboswitches in eubacteria sense the second messenger cyclic Di-GMP. *Science*, **321**, 411–413.

33. Pukatzki,S., Ma,A.T., Sturtevant,D., Krastins,B., Sarracino,D., Nelson,W.C., Heidelberg,J.F. and Mekalanos,J.J. (2006) Identification of a conserved bacterial protein secretion system in *Vibrio cholerae* using the *Science* host model system. *Proc. Natl. Acad. Sci. U. S. A.*, **103**, 1528–33.

34. Basler,M., Pilhofer,M., Henderson,G.P., Jensen,G.J. and Mekalanos,J.J. (2012) Type VI secretion requires a dynamic contractile phage tail-like structure. *Nature*, **483**, 182–186.

35. Van der Henst,C., Vanhove,A.S., Drebes Dörr,N.C., Stutzmann,S., Stoudmann,C., Clerc,S., Scrignari,T., Maclachlan,C., Knott,G. and Blokesch,M. (2018) Molecular insights into *Vibrio cholerae*'s intra-amoebal host-pathogen interactions. *Nat. Commun.*, **9**, 3460.

36. Unterweger,D., Kitaoka,M., Miyata,S.T., Bachmann,V., Brooks,T.M., Moloney,J., Sosa,O., Silva,D., Duran-Gonzalez,J., Provenzano,D., *et al.* (2012) Constitutive type VI secretion system expression gives *Vibrio cholerae* intra- and interspecific competitive advantages. *PLoS One*, **7**, e48320.

37. Bernardy,E.E., Turnsek,M.A., Wilson,S.K., Tarr,C.L. and Hammer,B.K. (2016) Diversity of clinical and environmental isolates of *Vibrio cholerae* in natural tansformation and contact-dependent bacterial

killing indicative of type VI secretion system activity. *Appl. Environ. Microbiol.*, **82**, 2833–2842.

38. Matthey,N., Stutzmann,S., Stoudmann,C., Guex,N., Iseli,C. and Blokesch,M. (2019) Neighbor predation linked to natural competence fosters the transfer of large genomic regions in *Vibrio cholerae*. *Elife*, **8**, e48212.

39. Matthey,N., Drebes Dörr,N.C. and Blokesch,M. (2018) Long-read-based genome sequences of pandemic and environmental *Vibrio cholerae* strains. *Microbiol. Resour. Announc.*, **7**.

40. Aldova,E., Laznickova,K., Stepankova,E. and Lietava,J. (1968) Isolation of nonagglutinable vibrios from an enteritis outbreak in Czechoslovakia. *J. Infect. Dis.*, **118**, 25–31.

41. Chun,J., Grim,C.J., Hasan,N.A., Lee,J.H., Choi,S.Y., Haley,B.J., Taviani,E., Jeon,Y.-S., Kim,D.W., Lee,J.-H., *et al.* (2009) Comparative genomics reveals mechanism for short-term and long-term clonal transitions in pandemic *Vibrio cholerae*. *Proc. Natl. Acad. Sci.*, **106**, 15442–15447.

42. Van der Henst,C., Vanhove,A.S., Drebes Dörr,N.C., Stutzmann,S., Stoudmann,C., Clerc,S., Scrignari,T., Maclachlan,C., Knott,G. and Blokesch,M. (2018) Molecular insights into *Vibrio cholerae*'s intra-amoebal host-pathogen interactions. *Nat. Commun.*, **9**, 3460.

43. Ben-Yaakov,R. and Salomon,D. (2019) The regulatory network of *Vibrio parahaemolyticus* type VI secretion system 1. *Environ. Microbiol.*, **21**, 2248–2260.

44. Kitaoka,M., Miyata,S.T., Brooks,T.M., Unterweger,D. and Pukatzki,S. (2011) VasH is a transcriptional regulator of the type VI secretion system functional in endemic and pandemic *Vibrio cholerae*. *J. Bacteriol.*, **193**, 6471–6482.

45. Drebes Dörr,N.C. and Blokesch,M. (2020) Interbacterial competition and anti-predatory behavior of environmental *Vibrio cholerae* strains. *Environ. Microbiol.*, **22**, 4485–4504.

46. Finkelstein,R.A., Boesman-Finkelsteint,M. and Holt,P. (1983) *Vibrio cholerae* hemagglutinin / lectin / protease hydrolyzes. *Proc. Natl. Acad. Sci.*, **80**, 1092–1095.

47. Scott,M.E., Dossani,Z.Y. and Sandkvist,M. (2001) Directed polar secretion of protease from single cells of *Vibrio cholerae* via the type II secretion pathway. *Proc. Natl. Acad. Sci. U. S. A.*, **98**, 13978–13983.

48. Sandkvist,M., Michel,L.O., Hough,L.P., Morales,V.M., Bagdasarian,M., Koomey,M., Dirita,V.J. and Bagdasarian,M. (1997) General secretion pathway (eps) genes required for toxin secretion and outer membrane biogenesis in *Vibrio cholerae*. *J. Bacteriol.*, **179**, 6994–7003.

49. Lin,L., Lezan,E., Schmidt,A. and Basler,M. (2019) Abundance of bacterial Type VI secretion system components measured by targeted proteomics. *Nat. Commun.*, **10**, 2584.

50. Seitz,P., Pezeshgi Modarres,H., Borgeaud,S., Bulushev,R.D., Steinbock,L.J., Radenovic,A., Dal Peraro,M. and Blokesch,M. (2014) ComEA is essential for the transfer of external DNA into the periplasm in naturally transformable *Vibrio cholerae* cells. *PLoS Genet.*, **10**, e1004066.

51. Joshi,A., Kostiuk,B., Rogers,A., Teschler,J., Pukatzki,S. and Yildiz,F.H. (2017) Rules of engagement: the type VI secretion system in *Vibrio cholerae*. *Trends Microbiol.*, **25**, 267–279.

52. Shrivastava,S. and Mande,S.S. (2008) Identification and functional characterization of gene components of type VI secretion system in bacterial genomes. *PLoS One*, **3**, e2955.

53. Zheng,J., Shin,O.S., Cameron,D.E. and Mekalanos,J.J. (2010) Quorum sensing and a global regulator TsrA control expression of type VI secretion and virulence in *Vibrio cholerae*. *Proc. Natl. Acad. Sci.*, **107**, 21128–21133.

54. Basler,M., Ho,B.T. and Mekalanos,J.J. (2013) Tit-for-tat: Type VI secretion system counterattack during bacterial cell-cell interactions. *Cell*, **152**, 884–894.

55. Cianfanelli,F.R., Monlezun,L. and Coulthurst,S.J. (2016) Aim, load, fire: the type VI secretion system, a bacterial nanoweapon. *Trends Microbiol.*, **24**, 51–62.

56. Mougous,J.D., Gifford,C.A., Ramsdell,T.L. and Mekalanos,J.J. (2007) Threonine phosphorylation post-translationally regulates protein secretion in *Pseudomonas aeruginosa*. *Nat. Cell Biol.*, **9**, 797–803.

57. Hsu,F., Schwarz,S. and Mougous,J.D. (2009) TagR promotes PpkA-catalysed type VI secretion activation in *Pseudomonas aeruginosa*. *Mol. Microbiol.*, **72**, 1111–1125.

58. Meibom,K.L., Li,X.B., Nielsen,A.T., Wu,C.-Y., Roseman,S. and Schoolnik,G.K. (2004) The *Vibrio cholerae* chitin utilization program. *Proc. Natl. Acad. Sci.*, **101**, 2524–2529.

59. Keene,O.N. (1995) The log transformation is special. *Stat. Med.*, **14**, 811–819.

60. Schindelin,J., Arganda-Carreras,I., Frise,E., Kaynig,V., Longair,M., Pietzsch,T., Preibisch,S., Rueden,C., Saalfeld,S., Schmid,B., *et al.* (2012) Fiji: an open-source platform for biological-image analysis. *Nat. Methods*, **9**, 676–682.

61. Lo Scrudato,M. and Blokesch,M. (2012) The regulatory network of natural competence and transformation of *Vibrio cholerae*. *PLoS Genet.*, **8**, e1002778.

62. Langmead,B. and Salzberg,S.L. (2012) Fast gapped-read alignment with Bowtie 2. *Nat. Methods*, **9**, 357–359.

63. Anders,S., Pyl,P.T. and Huber,W. (2015) HTSeq-A Python framework to work with high-throughput sequencing data. *Bioinformatics*, **31**, 166–169.

64. Love,M.I., Huber,W. and Anders,S. (2014) Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol.*, **15**.

65. Maaten Van Der,L. (2014) Accelerating t-SNE using tree-based algorithms. *J. Mach. Learn. Res.*, **15**, 3221–3245.

66. Yildiz,F.H. and Schoolnik,G.K. (1998) Role of rpoS in stress survival and virulence of *Vibrio cholerae*. *J. Bacteriol.*, **180**, 773–784.

67. Keymer,D.P., Miller,M.C., Schoolnik,G.K. and Boehm,A.B. (2007) Genomic and phenotypic diversity of coastal *Vibrio cholerae* strains is linked to environmental factors. *Appl. Environ. Microbiol.*, **73**, 3705–3714.

68. Miller,V.L., DiRita,V.J. and Mekalanos,J.J. (1989) Identification of toxS, a regulatory gene whose product enhances ToxR-mediated activation of the cholera toxin promoter. *J. Bacteriol.*, **171**, 1288–1293.

69. Pearson,G.D.N., Woods,A., Chiang,S.L. and Mekalanos,J.J. (1993) CTX genetic element encodes a site-specific recombination system and an intestinal colonization factor. *Proc. Natl. Acad. Sci. U. S. A.*, **90**, 3750–3754.

70. Yanisch-Perron,C., Vieira,J. and Messing,J. (1985) Improved M13 phage cloning vectors and host strains: nucleotide sequences of the M13mpl8 and pUC19 vectors. *Gene*, **33**, 103–119.

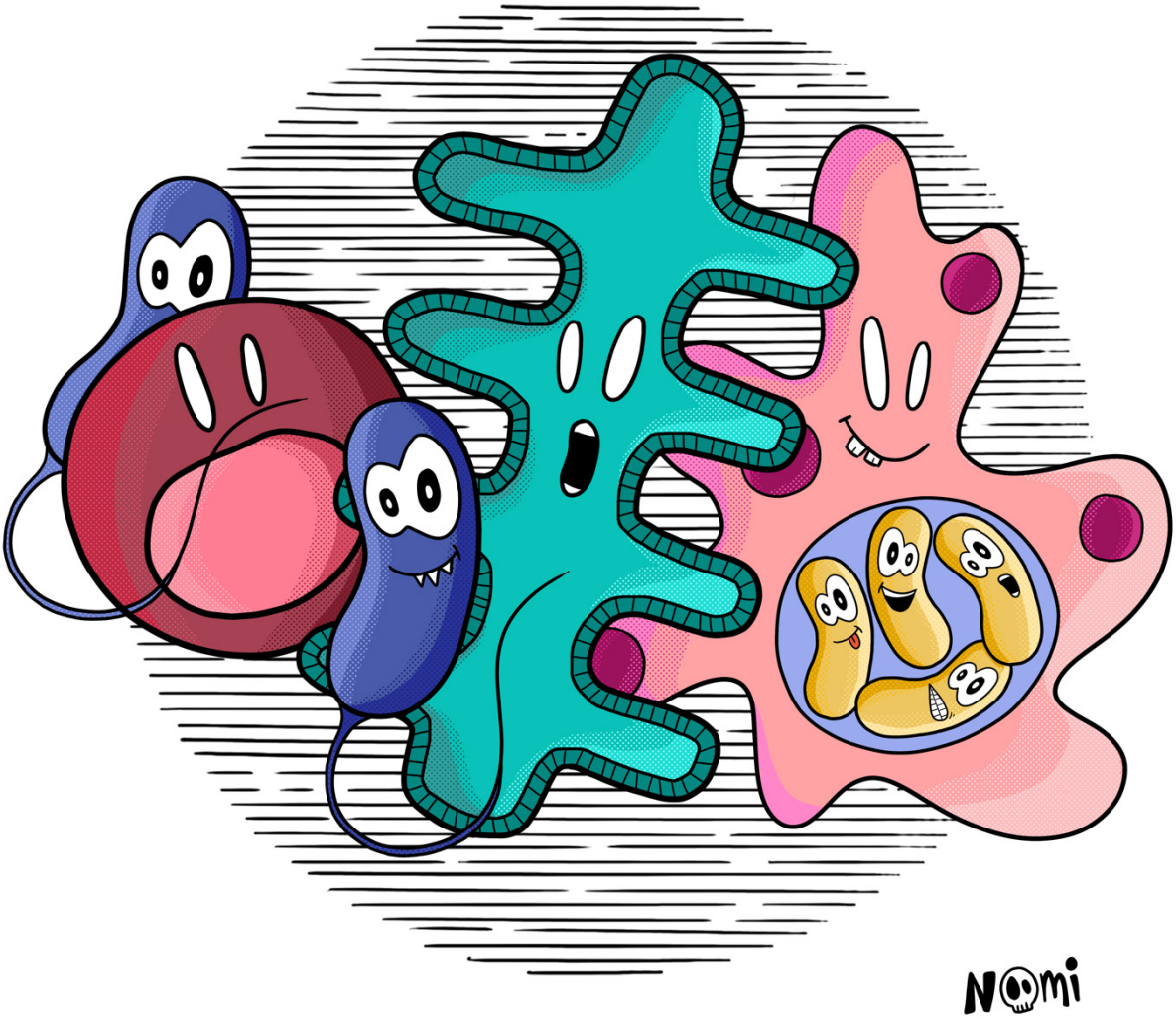
71. Simon,R., Preifer,U. and Phuler,A. (1983) A broad host range mobilization system for in vivo genetic enginering: transposon mutagenesis in gram negative bacteria. *Nat. Biotechnol.*, **1**, 784–791.

72. de Souza Silva,O. and Blokesch,M. (2010) Genetic manipulation of *Vibrio cholerae* by combining natural transformation with FLP recombination. *Plasmid*, **64**, 186–195.

73. Metzger,L.C., Matthey,N., Stoudmann,C., Collas,E.J. and Blokesch,M. (2019) Ecological implications of gene regulation by TfoX and TfoY among diverse *Vibrio* species. *Environ. Microbiol.*, **21**, 2231–2247.

74. Heidelberg,J.F., Elsen,J.A., Nelson,W.C., Clayton,R.A., Gwinn,M.L., Dodson,R.J., Haft,D.H., Hickey,E.K., Peterson,J.D., Umayam,L., *et al.* (2000) DNA sequence of both chromosomes of the cholera pathogen *Vibrio cholerae*. *Nature*, **406**, 477–483.

6 CONTRIBUTIONS TO OTHER PROJECTS



Drawing by Noémie Matthey, PhD

6.1 OVERVIEW

Over the course of my doctoral studies, I was lucky enough to get the opportunity to contribute to projects spearheaded by other lab members or in collaboration with other groups at EPFL. This chapter provides a very brief description of these contributions and some of the key findings achieved through my support.

6.2 INTERACTIONS OF *VIBRIO CHOLERAE* WITH *ACANTHAMOEBA CASTELLANII*

The data presented below were included in the publication:

Van der Henst C., Vanhove A.S., **Drebes Dörr N.C.,** Stutzmann S., Stoudmann C., Clerc S., Scrignari T., Maclachlan C., Knott G., Blokesch M. (2018) Molecular insights into *Vibrio cholerae*'s intra-amoebal host-pathogen interactions. ***Nat. Commun.,*** 9:3460. doi: 10.1038/s41467-018-05976-x.

As previously mentioned in this thesis, our lab has discovered the interaction of *V. cholerae* with the aquatic amoeba *A. castellanii* (1). In the second publication from our group involving this system, we uncovered the molecular mechanisms allowing the intra-amoebal lifestyle of *V. cholerae*. Briefly, *A. castellanii* are able to graze on *V. cholerae*, but part of the ingested bacteria resists intracellular degradation. A subset of food vacuoles containing non-degraded *V. cholerae* eventually fuses with the contractile vacuole (CV), which is the osmoregulatory organelle of these protozoa. The CV is then used by the bacteria as a niche for active replication, which even endures amoebal encystation. We demonstrated that proper regulation of certain *V. cholerae* accessory toxins, such as HapA protease, hemolysin and lecithinase, is imperative to allow completion of the infection cycle (2).

Furthermore, considering the relevance of the type VI secretion system (T6SS) in the interaction of *V. cholerae* with soil amoebae (3–5), we checked whether the T6SS was also involved in this host-pathogen interaction. Given that the intra-amoebal experiments had

been done with a pandemic *V. cholerae* strain (e.g., T6SS-silent under the tested conditions), we investigated the possible role of the T6SS using the toxigenic non-pandemic strain ATCC25872. To do so, I demonstrated the strain's constitutive T6SS activity using fluorescence microscopy (visualization of VipA-sfGFP structures) (Figure 6.1a) and in an agar plate-based killing assay in which *E. coli* served as prey (Figure 6.1b). This strain was then employed in *A. castellanii* infections, but the results suggested that the T6SS caused no harm to the host, as the infected amoebae displayed morphologies comparable to those infected by the pandemic control strain (Figure 6.1c and d).

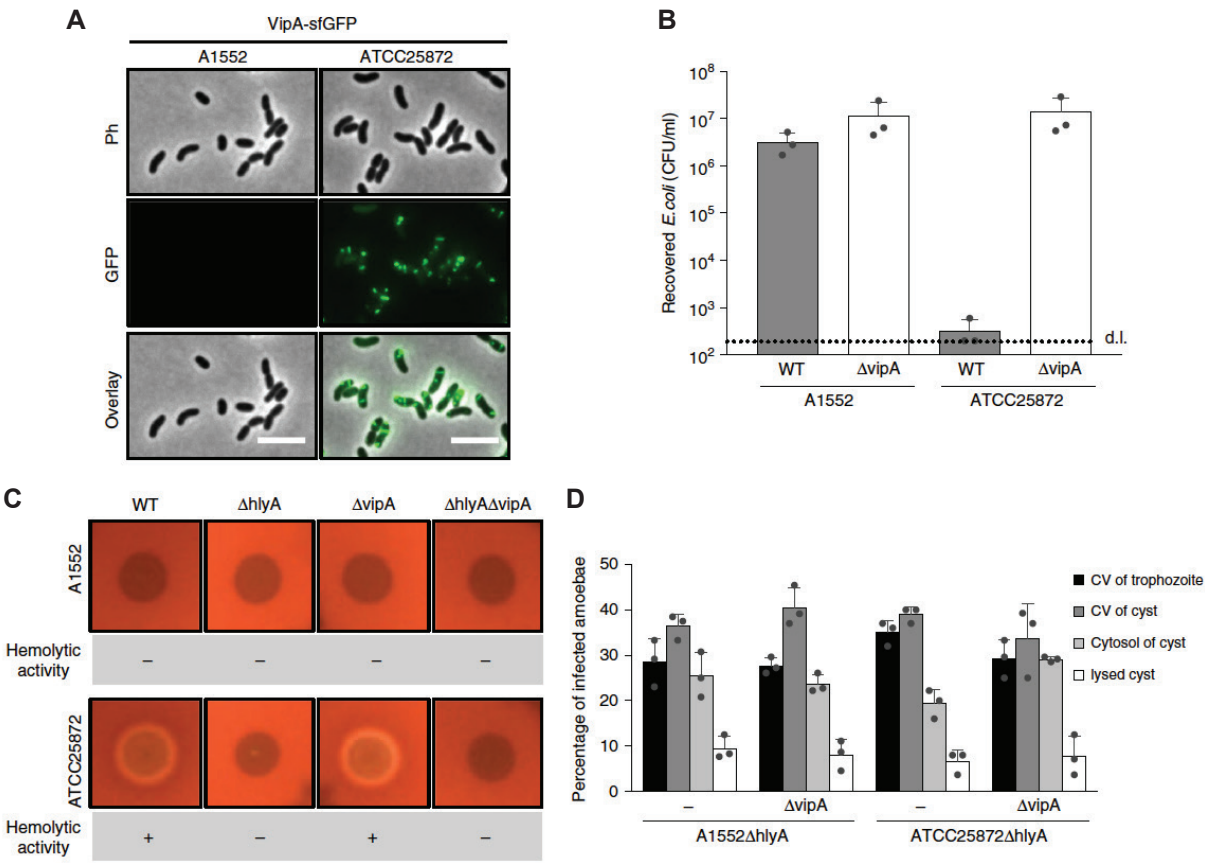


Figure 6.1 The T6SS of *V. cholerae* does not intoxicate *A. castellanii*.

(A-B) The T6SS of toxigenic non-pandemic strain ATCC25872 is constitutively active. **(A)** Fluorescence microscopy images of *V. cholerae* strains A1552 (O1 El Tor; pandemic) and ATCC25872 (O37; non-pandemic) carrying a translational fusion between the T6SS sheath protein VipA and sfGFP after growth for 3.5 h in LB. Cells are depicted in the phase contrast (Ph), green fluorescence (GFP) channel, and an overlay of both channels. Scale bar: 5 µm. **(B)** The constitutive T6SS activity of strain ATCC25872 was confirmed with bacterial killing assays using *E. coli* as prey. Numbers of surviving prey are depicted on the Y-axis (CFU/ml). Bar plots represent the average of three independent biological replicates (±SD). d.l., detection limit. **(C)** Hemolytic activity of strain ATCC25872 might mask T6SS effects. *V. cholerae* A1552 and ATCC25872 as well

as their Δ*hlyA*, Δ*vipA*, and Δ*hlyA*Δ*vipA* variants were tested for hemolytic activity on blood agar plates. **(D)** Mutant *V. cholerae* A1552 and ATCC25872 strains lacking hemolysin (Δ*hlyA*) or hemolysin and the T6SS core protein VipA (Δ*hlyA*Δ*vipA*) were assessed for their intra-amoebal localization at 20 h post-primary contact (p.p.c.). The considered compartments were the CV of trophozoites, the CV of cysts, and the cytosol of cysts before and after lysis. Values represent averages from three independent experiments (±s.d., as shown by the error bars). Figure reproduced from Fig. 4, panels c to f in reference (2) and legend adapted accordingly.

Another interesting aspect related to this study was the question of how conserved the intra-amoebal lifestyle was in different *V. cholerae* strains. I therefore constructed non-pandemic variant strains (ATCC25872 and Sa5Y) that lacked *hapA* and/or *hlyA* (as well as *vipA* alone or in combination with *hlyA*). Using these strains, we observed a high proportion of aberrantly-shaped amoebae when infected with these strains (Figures 6.1 and 6.2). Considering that non-pandemic strains constitutively secrete hemolysin, these results are consistent with what was observed for pandemic strains lacking the hemolysin-cleaving HapA protease (Figure 6.2).

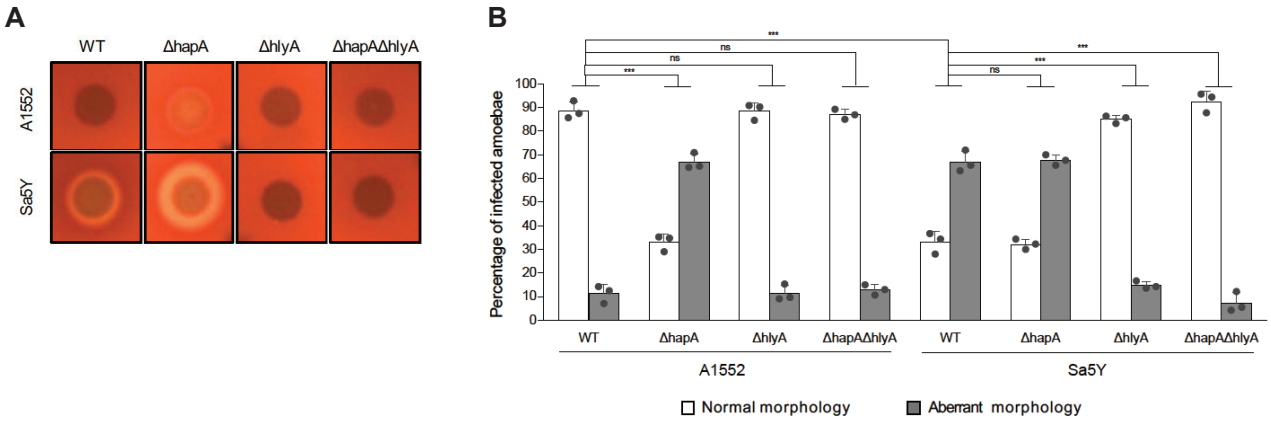


Figure 6.2 Amoebal hemolysin-dependent intoxication by environmental *V. cholerae*.

Pandemic *V. cholerae* strain A1552 and Californian environmental isolate Sa5Y (6) were tested for hemolytic activity on blood agar plates **(A)** and used for *A. castellanii* infections **(B)**, where amoebae with normal and aberrant morphologies were quantified 20 hours p.p.c. For each strain, the WT and their respective Δ*hapA*, Δ*hlyA* and Δ*hapA*Δ*hlyA* variants were tested. Values represent averages from three independent experiments (±s.d.). Statistics are based on one-way ANOVA with ***, $p \leq 0.001$; ns, $p > 0.05$. Figure reproduced from Fig. S5, panels a and c in reference (2) and legend adapted accordingly.

6.3 INTOXICATION OF *DICTYOSTELIUM DISCOIDEUM* BY *VIBRIO VULNIFICUS* IS IRON-DEPENDENT

At the end of 2018, our group hosted the visiting PhD student Carla Hernández-Cabanyero, coming from the laboratory of Prof. Carmen Amaro at the University of Valencia, Spain. The Amaro group has been working for many years on *V. vulnificus*, an important pathogen of humans and aquatic animals such as fish and oysters (7). When Carla visited our group, I trained her and collaborated with her to decipher any possible interactions between *V. vulnificus* and *D. discoideum* amoebae, given that I had already established the respective confocal microscopy-based approach (chapter 2). Intriguingly, we observed that *D. discoideum* cells exposed to exponentially growing *V. vulnificus* were completely unaffected (Figure 6.3a), in sharp contrast to the prompt amoebal cell rounding and intoxication exerted by *V. cholerae* (chapter 2). Nevertheless, the Amaro lab had demonstrated over the last years the relevance of environmental factors in the incidence and severity of vibriosis caused by *V. vulnificus*. Water temperature is particularly important for the onset of fish vibriosis, while iron availability is a major determinant in the severity of human vibriosis (7–9). Specifically, the disease in healthy humans is usually self-limiting, as iron - a very important factor for *V. vulnificus* virulence - is sequestered by the host's transferrin. In contrary, humans with chronic liver diseases and hemochromatosis, which leads to iron-overload in the serum, are highly susceptible to *V. vulnificus* septicemia making such iron overload a common risk factor (7, 10). We therefore tested the effect of iron supplementation (during growth and bacteria-amoebal co-cultures) in our system. Strikingly, under these conditions, the amoebae rapidly rounded up and developed the same intoxication phenotype that we observed when cells were exposed to *V. cholerae* (Figure 6.3b), demonstrating that *D. discoideum* serves as a good model to investigate *V. vulnificus* iron-dependent virulence. Notably, this finding might help us in the future to identify the toxic effector responsible for

D. discoideum intoxication by *V. cholerae*, as we could compare the transcriptomic data from our strains to that of *V. vulnificus* grown with or without iron (9, 11).

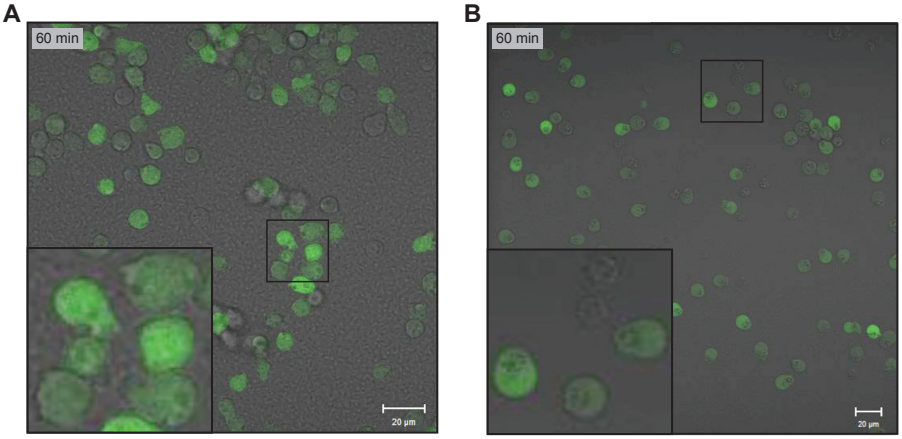


Figure 6.3 *D. discoideum* intoxication by *V. vulnificus* is iron-dependent.

D. discoideum cells were mixed with exponentially growing *V. vulnificus* strain R99 in the absence **(A)** or presence **(B)** of 1µM iron and imaged 60 min post primary contact (p.p.c). Lower left: magnification of the boxed area. Scale bars, 20 µm.

6.4 POTENTIAL OF MINI-INTESTINES FOR THE STUDY OF *V. CHOLERAE* INFECTION AND COLONIZATION

I also had the opportunity to actively participate in pilot experiments aiming at evaluating the potential use of mini-intestines/mini-guts in the study of host-pathogen interactions, in collaboration with the group of Prof. Matthias Lütolf (Laboratory of Stem Cell Bioengineering, EPFL). The Lütolf lab developed this tool as an alternative to three dimensional organoids, which usually result in tissues with a closed architecture, restricting their manipulation. The mini-guts are assembled with intestinal stem cells that develop inside a microchannel that mimics the gut's anatomical structure (with crypts and villi), and with an accessible lumen. An external and easy-to-use pumping system thereby allows perfusion of the mini-intestines allowing the removal of dead cells that are shed into the lumen (published ~ 2 years later by the Lütolf lab (12)).

Besides working on the development of properly homeostatic mini-intestines, the Lütolf lab was also interested in assessing the system's potential use for host-pathogen interactions. As *V. cholerae* in an important pathogen of the small intestine, we collaborated to develop a pilot infection system for *V. cholerae* and follow the bacterium's colonization using confocal microscopy-based visualization of these mini-guts. We performed many experiments and optimizations over a period of approximately two months. However, I will solely summarize three interesting observations that we made throughout this study.

When two identical wild-type (WT) pandemic El Tor *V. cholerae* strains, one tagged with mCherry and the other with GFP, were mixed together and inoculated simultaneously inside the mini-intestine, there was no clear "advantage" of one strain over the other, as expected. Both strains were able to enter the mini-gut, reach the intestinal crypts and replicate over time, thereby forming small colonies (Figure 6.4).

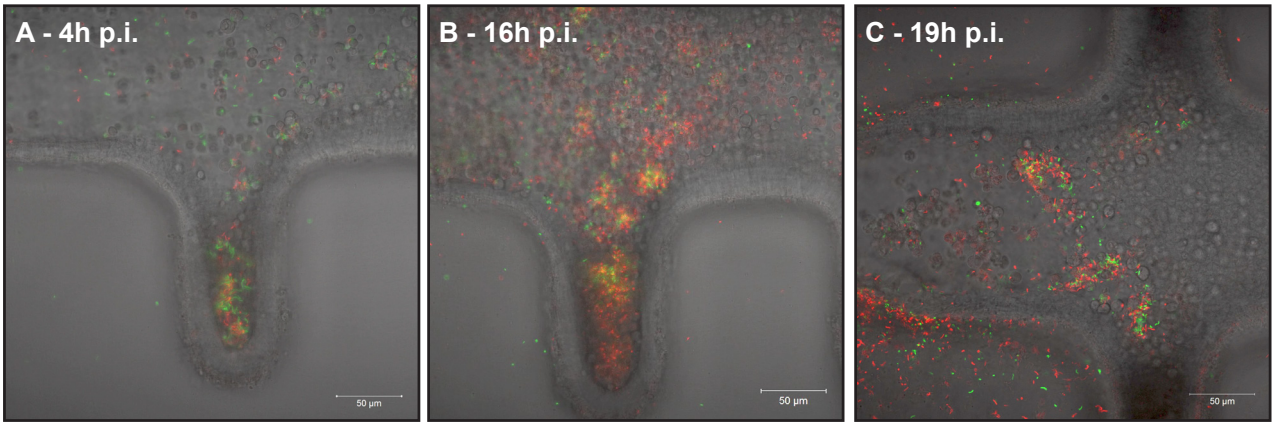


Figure 6.4 *V. cholerae* is capable of colonizing the crypts of mini-intestine.

Two identical wild-type *V. cholerae* strains (pandemic El Tor A1552), one producing GFP and the other mCherry, were grown for 3 h, washed and mixed at a 1:1 ratio, and then inoculated into the mini-gut after 1:10 dilution in Advanced F-12 DMEM medium supplemented with growth factors. The inoculated mini-gut was kept at 37°C in a CO_2 incubator and then imaged 4 h **(A)**; 16 h **(B)** and 19 h **(C)** post inoculation (p.i.). Scale bars, 50 µm.

Interestingly, however, when we mixed and inoculated a WT mCherry-tagged pandemic strain with a GFP-tagged variant lacking the virulence transcriptional regulator

ToxT, we observed a clear separation of the two strains inside the mini-intestine. Specifically, while the WT strain was mainly restricted to inlet-side of the mini-gut and to the first pair of crypts, the Δ*toxT* strain localized primarily downstream of the WT and specifically to the second pair of crypts and the outlet of the device (Figure 6.5). Of note, the intestine portions colonized by the WT strain seemed to be collapsing, while the parts colonized by the mutant still looked healthy and intact.

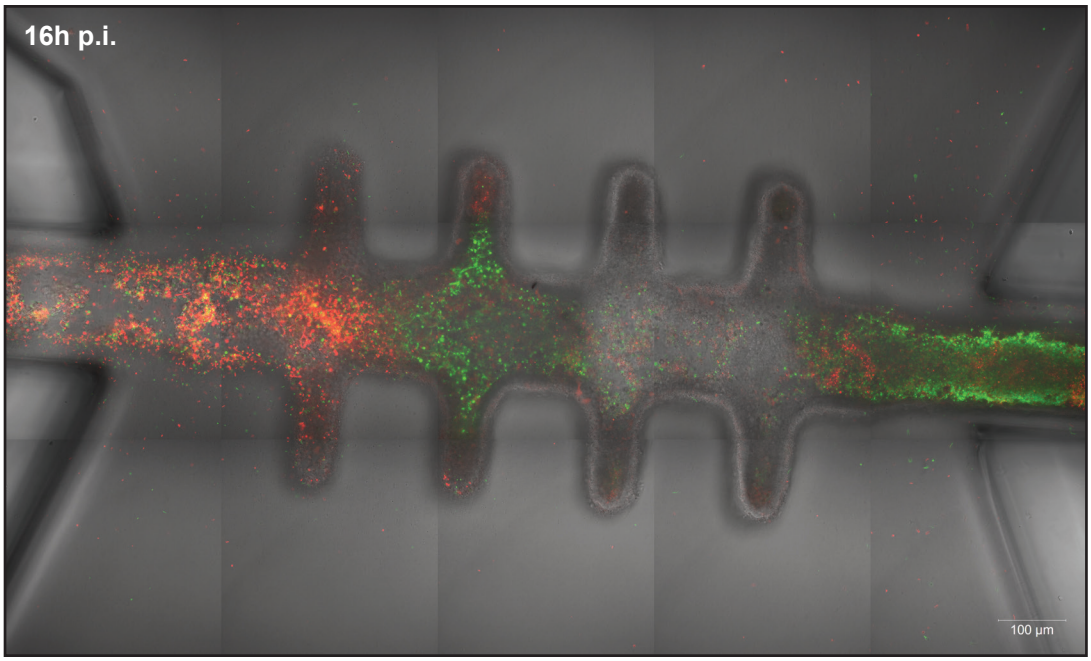


Figure 6.5 Virulence factors seem to be involved in mini-gut colonization.

A WT pandemic *V. cholerae* strain producing mCherry and a Δ*toxT* variant producing GFP were grown for 3 h, washed and mixed at a 1:1 ratio, and then inoculated into the mini-gut after 1:10 dilution in Advanced F-12 DMEM medium supplemented with growth factors. The inoculated mini-gut was kept at 37°C in a CO_2 incubator and then imaged 16 h post inoculation (p.i.). Scale bar, 100 µm.

In pandemic *V. cholerae* strains, ToxT activates transcription of the *tcp* genes, which encode the toxin-coregulated pilus (TCP). TCP is the primary colonization factor of *V. cholerae*, which fosters auto-aggregation and adherence to the epithelium of the small intestine (13, 14). To further investigate the separation of the two strains observed in Figure 6.5, we next tested a mCherry-tagged strain that carried a previously described *tcpA* (major

TCP pilin) transcriptional reporter construct (15). 16 h post-inoculation (p.i.), we observed that the bacteria had spread throughout the mini-gut (Figure 6.6a). A strong GFP signal derived from *tcpA* expression started to appear after around 27 h p.i. (Figure 6.6b), when cells formed aggregates in the lumen, which spread over time (Figure 6.6c). Interestingly, at around 72 h p.i., *V. cholerae* cells became completely round (Figure 6.6d). It will be interesting to test the reproducibility of these findings and to decipher their potential biological relevance. Notably, our group recently discovered a cell rounding mechanism in *V. cholerae* and future studies could show if this mechanism is induced in the mini-gut system.

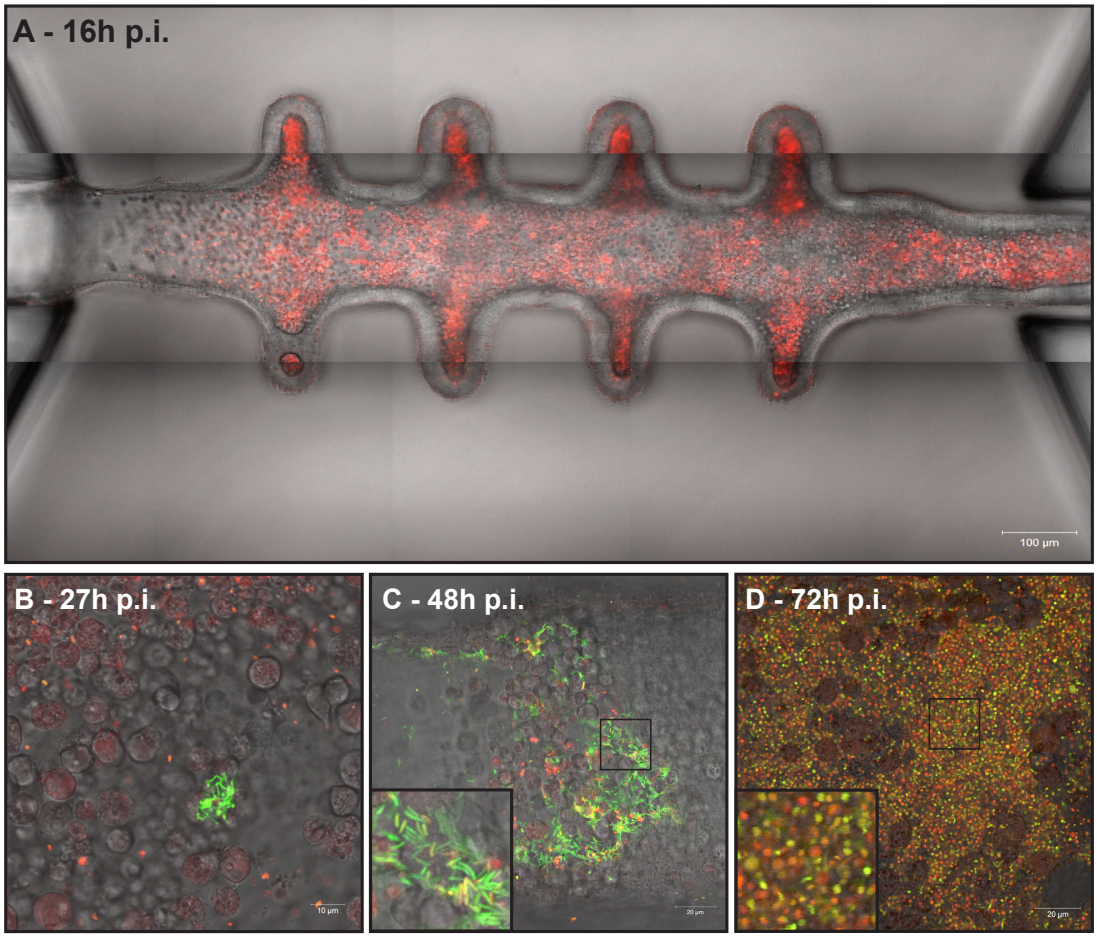


Figure 6.6 Toxin co-regulated pilus (TCP) induction and *V. cholerae* cell rounding inside mini-intestine.

A *V. cholerae* strain constitutively expressing mCherry and carrying a transcriptional reporter for *tcpA* expression [promoter fused to GFP(ASV)] was grown for 3 h, washed and then inoculated into the mini-gut after 1:10 dilution in Advanced F-12 DMEM medium supplemented with growth factors. The inoculated mini-gut was kept at 37°C in a CO_2 incubator and then imaged 16 h **(A)**; 27 h **(B)**; 48 h **(C)** and 72 h **(D)** post

inoculation (p.i.). Lower left: magnification of the boxed area. Scale bars 100 µm (A); 10 µm (B) and 20 µm (C and D).

Importantly, however, what we observed in all of our experiments was that the mini-intestines were “leaky”, especially at the ends of the tube. This leakage allowed *V. cholerae* cells to reach the highly nutritious hydrogel that surrounds the mini-guts with the purpose to nourish the stem cells. Consequently, extensive bacterial growth occurred in this environment (as exemplified in (Figure 6.4C) which, ultimately, led to the collapse of the mini-intestines. Even though the Lütolf lab employed different strategies to circumvent the problem of leakage and bacterial replication in the hydrogel, we could not reach a point in which experiments could be reliably replicated beyond the 12 hours p.i. stage. The project was therefore discontinued from the side of our collaborators, who preferred to assay the mini-guts with other pathogens, namely *Cryptosporidium parvum* (12).

6.5 REFERENCES

1. Van Der Henst,C., Scrignari,T., Maclachlan,C. and Blokesch,M. (2016) An intracellular replication niche for *Vibrio cholerae* in the amoeba *Acanthamoeba castellanii*. *ISME J.*, **10**, 897–910.

2. Van der Henst,C., Vanhove,A.S., Drebes Dörr,N.C., Stutzmann,S., Stoudmann,C., Clerc,S., Scrignari,T., Maclachlan,C., Knott,G. and Blokesch,M. (2018) Molecular insights into *Vibrio cholerae*'s intra-amoebal host-pathogen interactions. *Nat. Commun.*, **9**, 3460.

3. Pukatzki,S., Ma,A.T., Sturtevant,D., Krastins,B., Sarracino,D., Nelson,W.C., Heidelberg,J.F. and Mekalanos,J.J. (2006) Identification of a conserved bacterial protein secretion system in *Vibrio cholerae* using the *Dictyostelium* host model system. *Proc. Natl. Acad. Sci. U. S. A.*, **103**, 1528–33.

4. Unterweger,D., Kitaoka,M., Miyata,S.T., Bachmann,V., Brooks,T.M., Moloney,J., Sosa,O., Silva,D., Duran-Gonzalez,J., Provenzano,D., *et al.* (2012) Constitutive type VI secretion system expression gives *Vibrio cholerae* intra- and interspecific competitive advantages. *PLoS One*, **7**, e48320.

5. Drebes Dörr,N.C. and Blokesch,M. (2020) Interbacterial competition and anti-predatory behavior of environmental *Vibrio cholerae* strains. *Environ. Microbiol.*, **22**, 4485–4504.

6. Keymer,D.P., Miller,M.C., Schoolnik,G.K. and Boehm,A.B. (2007) Genomic and phenotypic diversity of coastal *Vibrio cholerae* strains is linked to environmental factors. *Appl. Environ. Microbiol.*, **73**, 3705–3714.

7. Hernández-Cabanyero,C. and Amaro,C. (2020) Phylogeny and life cycle of the zoonotic pathogen *Vibrio vulnificus*. *Environ. Microbiol.*, **22**, 4133–4148.

8. Hernández-Cabanyero,C., Sanjuán,E., Fouz,B., Pajuelo,D., Vallejos-Vidal,E., Reyes-López,F.E. and Amaro,C. (2020) The effect of the environmental temperature on the adaptation to host in the zoonotic pathogen *Vibrio vulnificus*. *Front. Microbiol.*, **11**.

9. Pajuelo,D., Hernández-Cabanyero,C., Sanjuan,E., Lee,C. Te, Silva-Hernández,F.X., Hor,L.I., Mackenzie,S. and Amaro,C. (2016) Iron and fur in the life cycle of the zoonotic pathogen *Vibrio vulnificus*. *Environ. Microbiol.*, **18**, 4005–4022.

10. Horseman,M.A. and Surani,S. (2011) A comprehensive review of *Vibrio vulnificus*: an important cause of severe sepsis and skin and soft-tissue infection. *Int. J. Infect. Dis.*, **15**, e157–e166.

11. Hernández-Cabanyero,C., Lee,C. Te, Tolosa-Enguis,V., Sanjuán,E., Pajuelo,D., Reyes-López,F., Tort,L. and Amaro,C. (2019) Adaptation to host in *Vibrio vulnificus*, a zoonotic pathogen that causes septicemia in fish and humans. *Environ. Microbiol.*, **21**, 3118–3139.

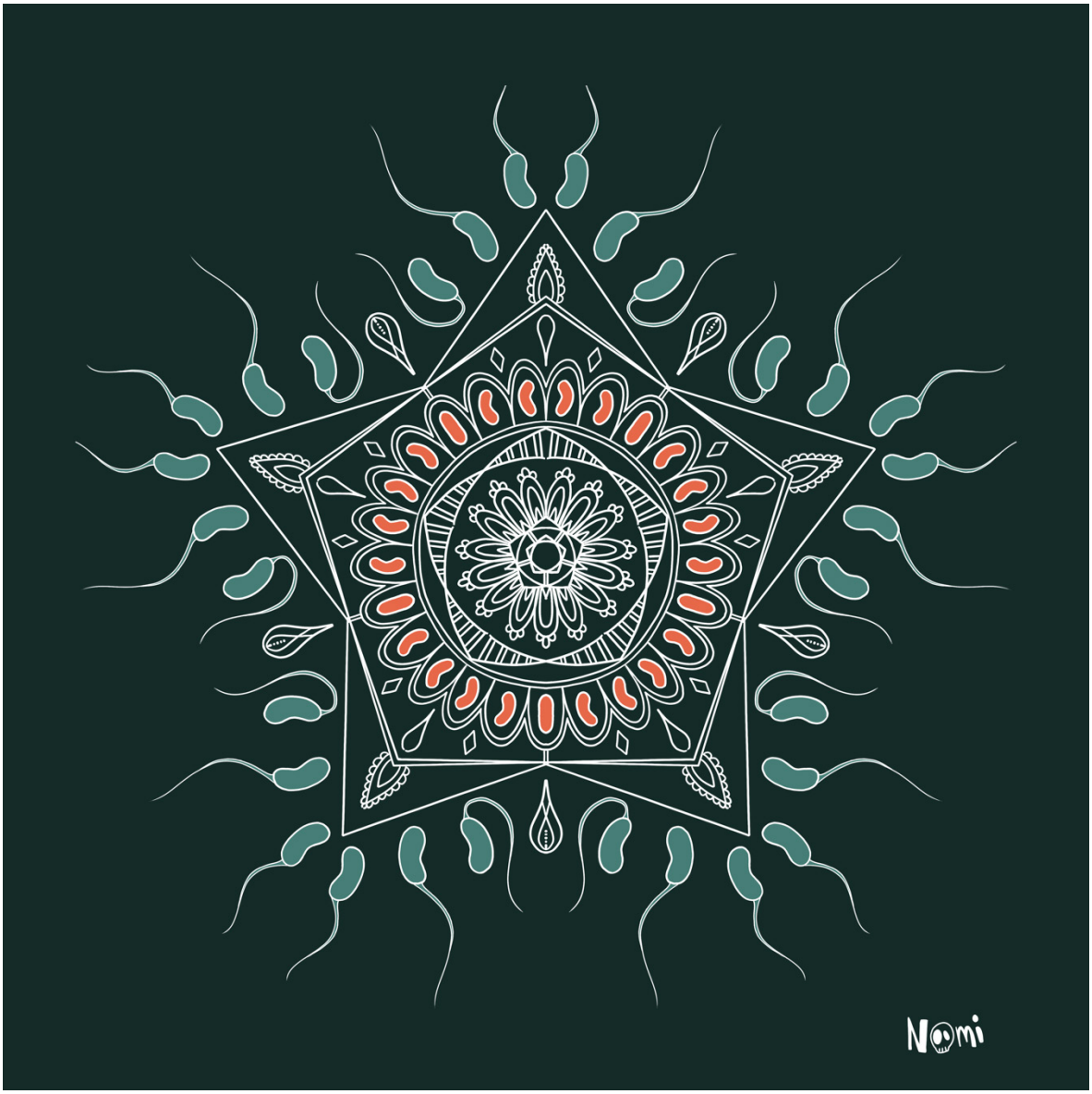
12. Nikolaev,M., Mitrofanova,O., Broguiere,N., Geraldo,S., Dutta,D., Tabata,Y., Elci,B., Brandenberg,N., Kolotuev,I., Gjorevski,N., *et al.* (2020) Homeostatic mini-intestines through scaffold-guided organoid morphogenesis. *Nature*, **585**, 574–578.

13. Taylor,R.K., Miller,V.L., Furlong,D.B. and Mekalanos,J.J. (1987) Use of *phoA* gene fusions to identify a pilus colonization factor coordinately regulated with cholera toxin. *Proc. Natl. Acad. Sci.*, **84**, 2833–2837.

14. Kirn,T.J., Lafferty,M.J., Sandoe,C.M.P. and Taylor,R.K. (2000) Delineation of pilin domains required for bacterial association into microcolonies and intestinal colonization by *Vibrio cholerae*. *Mol. Microbiol.*, **35**, 896–910.

15. Nielsen,A.T., Dolganov,N.A., Rasmussen,T., Otto,G., Miller,M.C., Felt,S.A., Torreilles,S. and Schoolnik,G.K. (2010) A bistable switch and anatomical site control *Vibrio cholerae* virulence gene expression in the intestine. *PLoS Pathog.*, **6**.

7 CONCLUSION AND PERSPECTIVES



Drawing by Noémie Matthey, PhD

Cholera pandemics have afflicted mankind for centuries (1–3), and annual case numbers revolve around 4 million people still nowadays (4). Nonetheless, even though research has uncovered an outstanding level of detail into many biological aspects of its causative agent, *V. cholerae*, we still don't fully understand the complete set of factors that allowed the emergence and success of the pandemic lineage. Similar to their diverse and non-pathogenic environmental counterparts, *V. cholerae* pandemic strains can be found in the aquatic environment. In this habitat, the bacteria are commonly found associated with chitinous surfaces from zooplankton. Chitin serves as a nutrition source but is also responsible for the induction of important biological programs, such as T6SS-mediated neighbor killing and natural competence (5–7). Furthermore, much of cholera's epidemiology is based on environmental factors such as annual monsoons and zooplankton blooms (due to warmer waters), especially in the endemic region of Bangladesh (8–10). Therefore, understanding what makes the pandemic lineage special might help us better understand pandemicity evolution (1, 9, 11, 12). In that sense, this thesis aimed at investigating potentially relevant aspects of *V. cholerae*'s biology in a comparative framework.

Avoiding protozoan predation is crucial for the survival of bacteria in the aquatic environment (13, 14). We therefore explored possible strategies used by different *V. cholerae* strains when facing a model eukaryotic grazer, *D. discoideum*. Importantly, these amoebae were used in the experiments that first led to the discovery of the T6SS in *V. cholerae* (15). As previous studies used grazing assays on solid surfaces, we were interested in visualizing such interactions under liquid conditions and to follow them by time-lapse microscopy. Accordingly, we established confocal laser scanning microscopy-based time-lapse imaging experiments in which we observed quick intoxication of the amoebae by *V. cholerae*. Our results therefore confirmed that these amoebae can serve as a powerful

tool to screen for *Vibrio* virulence factors. Furthermore, despite the fact that we were not able to determine the *V. cholerae* toxic effector responsible for amoebal intoxication up to now, we obtained preliminary data that will guide future directions of this project. For example, we observed environmental strains that were able to intoxicate the amoebae even when grown to high cell density (HCD), in contrast to other environmental strains and the pandemic lineage. Additionally, we observed that *V. vulnificus* only intoxicated the amoebae when grown with iron, which induces virulence in this pathogen. Going forward, we will investigate thoroughly the transcriptomic data to identify the causative factor for the observed intoxication.

Additional work developed during this thesis in the realm of host-pathogen interactions involved two collaborative projects. First, we demonstrated that the T6SS is not involved in the intra-amoebal lifestyle of *V. cholerae* when interacting with the aquatic amoeba *A. castellanii*. Instead, these bacteria finetune the expression and secretion of a set of extracellular enzymes in order to complete their intra-contractile vacuole cycle, which allows prolific replication and the release of bacteria back into the environment (16). Secondly, we performed pilot experiments with the Lütolf laboratory (EPFL) using mini-intestines (17) to study *V. cholerae* colonization. Even though the mini-gut setup was not fully functional yet for studies with highly replicative bacteria such as *V. cholerae*, we were able to observe important phenotypes, which would be very interesting to investigate further in the future.

V. cholerae strains from the 7th pandemic lineage have a very similar genomic organization and content, and much of their evolutionary trajectory has been established by comparative genomics and a SNP-based molecular clock (18). Environmental isolates, on the other hand, which represent most of *V. cholerae*'s diversity, have been far less studied than the pandemic lineage. We therefore took advantage of our long-read-based whole

genome sequencing data that we derived from a collection of environmental strains, and thoroughly characterized the mobilome present in these strains. Contrary to the high conservation of genomic islands found in pandemic strains, we observed that these environmental isolates were extremely diverse. We uncovered elements carrying genes that encoded factors that we classified into five main categories: (i) anti-phage defense; (ii) bacterial antagonism; (iii) cell appendages; (iv) metabolism; and (v) interaction with host. Furthermore, through expression profiling, we demonstrated that most of these islands were transcriptionally active; future investigations are therefore planned to experimentally validate their functionality. Importantly, the PacBio technology-based sequencing data also provided us with information regarding the DNA modification landscape of these strains and a comparison group of pandemic strains. Here again the pandemic strains presented a highly conserved landscape. Environmental strains, on the other hand, displayed different and sometimes strain-specific DNA modifications patterns that were, in part, spatially localized. Future studies will identify the modifying enzymes of these strain-specific DNA marks (most likely localized on genomic islands) and investigate the biological relevance of the different DNA modifications. This knowledge might provide additional clues as to how pandemic and environmental strains regulate gene expression and which areas of their genomes are preferentially kept active or silenced.

Competition with other bacteria (19, 20) and protozoan predation (13, 14) are two of the major challenges faced by bacteria in the aquatic environment. In that sense, we were interested in assessing and comparing the potential of environmental or pandemic *V. cholerae* strains when encountering these threats. Specifically, we investigated the role of two minor virulence factors, the pore-forming toxin hemolysin and the T6SS. We demonstrated that all environmental strains kept the T6SS active under the tested (laboratory) conditions, as had been shown previously in other studies (21, 22). Constitutive

T6SS activity is in sharp contrast to what is observed in pandemic strains, which keep their T6SS under tight regulation (5, 23, 24). Interestingly, even though all environmental strains used their T6SS against bacteria, only a subset of them used it to avoid grazing by *D. discoideum* when confronted to each other on a solid surface. This phenotype was caused by an actin-crosslinking domain T6SS effector. Furthermore, a comprehensive *in silico* characterization uncovered a diverse array of T6SS effector/ immunity pairs carried by these environmental isolates. Consistently, pairwise competition experiments demonstrated a very low level of compatibility between these strains. It is reasonable to assume that constitutive T6SS activity and the effector/immunity diversity might provide these strains with fitness advantages in their natural habitat. Conversely, it might have been beneficial for pandemic strains to evolve a tight regulation of these molecular weapons, potentially in a disease context.

Finally, we tackled the phenomenon of T6SS constitutive activity in non-pandemic *V. cholerae* strains. As previously explained, T6SS expression is tightly controlled in pandemic strains, where significant activation can be induced by two independent routes controlled by one of two regulators, TfoX or TfoY (5, 23, 24). Conversely, metabolically costly constitutive T6SS production in non-pandemic isolates is not understood, even though the phenomenon has been observed since the T6SS discovery in 2006 (15). To tackle the underlying regulatory circuits, we developed a hybrid strain library, which resulted in a collection of 800 clones with a pandemic strain-based genetic backbone and containing different portions of the non-pandemic strain's genome. These hybrid clones were tested for their T6SS activity, which allowed us to ultimately identify a causative nucleotide (nt45) in an intergenic region of the T6SS large cluster. While the 7th pandemic lineage contains a 'G' nucleotide in this position, all non-pandemic strains harbor a 'T' single nucleotide polymorphism, which maintains the T6SS constitutively active. After extensive characterization of the nt45-linked

activity, we proposed a preliminary mode-of-action model in which we suggest that the regulation works through a positive feedback loop involving a still unidentified regulator encoded in the T6SS large cluster. Current efforts in this project involve experiments to identify this regulator and to ultimately decipher the regulatory mechanism responsible for the nt45-derived T6SS regulation.

In summary, this thesis provides exciting and novel findings in the realm of *V. cholerae* evolution. Importantly, most of our results are based on the molecular treasure box of the environmental strains, demonstrating how important it is to study non-pandemic isolates. Our results contribute to a better understanding of *V. cholerae*'s evolution from an innocuous inhabitant of aquatic environments to a very relevant, pandemic-causing, human pathogen.

7.1 REFERENCES

1. Colwell,R.R. (1996) Global climate and infectious disease: the cholera paradigm. *Science.*, **274**, 2025–2031.

2. Clemens,J.D., Nair,G.B., Ahmed,T., Qadri,F. and Holmgren,J. (2017) Cholera. *Lancet*, **390**, 1539–1549.

3. S.M.,F., M.J.,A. and J.J.,M. (1998) Epidemiology, genetics, and ecology of toxigenic *Vibrio cholerae*. *Microbiol. Mol. Biol. Rev.*, **62**, 1301–1314.

4. Ali,M., Nelson,A.R., Lopez,A.L. and Sack,D.A. (2015) Updated global burden of cholera in endemic countries. *PLoS Negl. Trop. Dis.*, **9**, e0003832.

5. Borgeaud,S., Metzger,L.C., Scrignari,T. and Blokesch,M. (2015) The type VI secretion system of *Vibrio cholerae* fosters horizontal gene transfer . *Science.*, **347**, 63–67.

6. Meibom,K.L., Blokesch,M., Dolganov,N.A., Wu,C.Y. and Schoolnik,G.K. (2005) Chitin induces natural competence in *Vibrio cholerae*. *Science.*, **310**, 1824–1827.

7. Blokesch,M. (2014) The lifestyle of *Vibrio cholerae* fosters gene transfers. *Microbe*, **9**, 64–70.

8. Emch,M., Feldacker,C., Sirajul,S.M. and Ali,M. (2008) Seasonality of cholera from 1974 to 2005: A review of global patterns. *Int. J. Health Geogr.*, **7**.

9. Lipp,E.K., Huq,A. and Colwell,R.R. (2002) Effects of global climate on infectious disease: the cholera model. *Clin. Microbiol. Rev.*, **15**, 757–770.

10. Lobitz,B., Beck,L., Huq,A., Wood,B., Fuchs,G., Faruque,A.S.G. and Colwell,R. (2000) Climate and infectious disease: use of remote sensing for detection of *Vibrio cholerae* by indirect measurement. *Proc. Natl. Acad. Sci.*, **97**, 1438–1443.

11. Martínez,J.L. (2013) Bacterial pathogens: from natural ecosystems to human hosts. *Environ. Microbiol.*, **15**, 325–333.

12. Alam,M., Kasan,N.A., Sadique,A., Bhuiyan,N.A., Ahmed,K.U., Nusrin,S., Nair,G.B., Siddique,A.K., Sack,R.B., Sack,D.A., *et al.* (2006) Seasonal cholera caused by *Vibrio cholerae* serogroups O1 and O139 in the coastal aquatic environment of Bangladesh. *Appl. Environ. Microbiol.*, **72**, 4096–4104.

13. Matz,C. and Kjelleberg,S. (2005) Off the hook - How bacteria survive protozoan grazing. *Trends Microbiol.*, **13**, 302–307.

14. Brüssow,H. (2007) Bacteria between protists and phages: from antipredation strategies to the evolution of pathogenicity. *Mol. Microbiol.*, **65**, 583–589.

15. Pukatzki,S., Ma,A.T., Sturtevant,D., Krastins,B., Sarracino,D., Nelson,W.C., Heidelberg,J.F. and Mekalanos,J.J. (2006) Identification of a conserved bacterial protein secretion system in *Vibrio cholerae* using the *Dictyostelium* host model system. *Proc. Natl. Acad. Sci. U. S. A.*, **103**, 1528–33.

16. Van der Henst,C., Vanhove,A.S., Drebes Dörr,N.C., Stutzmann,S., Stoudmann,C., Clerc,S., Scrignari,T., Maclachlan,C., Knott,G. and Blokesch,M. (2018) Molecular insights into *Vibrio cholerae*'s intra-amoebal host-pathogen interactions. *Nat. Commun.*, **9**, 3460.

17. Nikolaev,M., Mitrofanova,O., Broguiere,N., Geraldo,S., Dutta,D., Tabata,Y., Elci,B., Brandenberg,N., Kolotuev,I., Gjorevski,N., *et al.* (2020) Homeostatic mini-intestines through scaffold-guided organoid morphogenesis. *Nature*, **585**, 574–578.

18. Mutreja,A., Kim,D.W., Thomson,N.R., Connor,T.R., Lee,J.H., Kariuki,S., Croucher,N.J., Choi,S.Y., Harris,S.R., Lebens,M., *et al.* (2011) Evidence for several waves of global transmission in the seventh cholera pandemic. *Nature*, **477**, 462–465.

19. Granato,E.T., Meiller-Legrand,T.A. and Foster,K.R. (2019) The evolution and ecology of bacterial warfare. *Curr. Biol.*, **29**, R521–R537.

20. Peterson,S.B., Bertolli,S.K. and Mougous,J.D. (2020) The central role of interbacterial antagonism in bacterial fife. *Curr. Biol.*, **30**, R1203–R1214.

21. Unterweger,D., Kitaoka,M., Miyata,S.T., Bachmann,V., Brooks,T.M., Moloney,J., Sosa,O., Silva,D., Duran-Gonzalez,J., Provenzano,D., *et al.* (2012) Constitutive type VI secretion system expression gives *Vibrio cholerae* intra- and interspecific competitive advantages. *PLoS One*, **7**, e48320.

22. Bernardy,E.E., Turnsek,M.A., Wilson,S.K., Tarr,C.L. and Hammer,B.K. (2016) Diversity of clinical and environmental isolates of *Vibrio cholerae* in natural transformation and contact-dependent bacterial killing indicative of type VI secretion system activity. *Appl. Environ. Microbiol.*, **82**, 2833–2842.

23. Metzger,L.C., Stutzmann,S., Scrignari,T., Van der Henst,C., Matthey,N. and Blokesch,M. (2016) Independent regulation of type VI secretion in *Vibrio cholerae* by TfoX and TfoY. *Cell Rep.*, **15**, 951–958.

24. Jaskólska,M., Stutzmann,S., Stoudmann,C. and Blokesch,M. (2018) QstR-dependent regulation of natural competence and type VI secretion in *Vibrio cholerae*. *Nucleic Acids Res.*, **46**, 10619–10634.

ACKNOWLEDGEMENTS

If I look back at my five years of PhD, I cannot help having a big smile on my face. What a journey this has been! I first came to EPFL eleven years ago for the Life Sciences Summer Research Program, during my Bachelor's. Who knew that six years later I would finally be able to participate in the PhD selection at this very place and be matched with my favorite lab from the very beginning? Even though I had spent those six years dreaming about coming back to Switzerland, I never imagined that this experience would change me for better in so many ways, and that it would have such a huge impact in my life. I'm honored, humbled and most of all thankful for having had the chance to experience my PhD at this place in this exact moment in time. By now you can already tell that this will be a long list of acknowledgments… bear with me if you have the patience, but this needs to be long.

I need to start by thanking my PhD supervisor, Prof. Melanie Blokesch. I still pinch myself when I think about how lucky I am by having done my PhD with you. I came here so "green", without any real experience with Microbiology - but so very eager to learn. You were incredibly supportive in every step of the way, teaching me, guiding me but also trusting me. Besides all the Science that I've learned from you, you helped me understand what it takes to have a career in this field. During those moments that I thought I wasn't good enough, in which I doubted myself, you were there, with your door or e-mail always open, with words of encouragement but always decorated with a big dose of reality. That was priceless for my growth. Thank you for your incredible passion for Science, for your excitement and big smile when I came with new results, for the animated discussions and plans we did together over these years. That was so much fun.

Marisa, our superstar secretary, also known as “angel” or “life savior”, thank you for always having our back, for your support and emails with a smiley face. You found a place for Matheus and I to live here in Switzerland when we first arrived from Brazil, you picked us up at the metro station at 11pm with a big smile – I couldn’t believe we were so lucky! You were so caring, lovely and important to me during these years, I really cannot thank you enough!

A huge thank you to those in the lab who were there at my very beginning and who helped me find my way around the benches. Charles, you were invaluable! I wish I knew back then how much you were teaching me, when you showed me how to be focused, how to have clear questions to answer, how to take life without too much stress, how to laugh when things didn’t work, and how to be so incredibly ethical and trustworthy. Lisa, you showed me the power of commitment, of hard work, of being proactive and of community spirit. Thank you for always caring so much and for offering a chocolate when you thought I was overwhelmed (which can be quite often at the very beginning of a PhD). Noémie, you were my first example of what a PhD student should be like, and I have learned so very much from you! Your relentless dedication to have things done perfectly was so inspiring. Thank you for also teaching me dance moves and for sharing so many delicious recipes! And a huge thank you for all the amazing drawings you did for my thesis. You are incredibly talented, and I wish you get more and more recognition for your work with Science Communication. Sandrine (Stutzmann), our machine-technician, my dear friend and supporter who helped me in so many ways during these years! I know you did many experiments for me which you hated (hello supernantant fraction extraction) and for that I apologize. But I thank you for always being there, for teaching me how to speak French (which required patience!), for being so attentive and caring, for checking on me and offering a hand or your shoulder anytime I needed. David and Milena, you were real scientific role

models for me, with your energy to tackle questions in so many fronts, your dedication, your perfectionism, your clear and beautiful lab meetings and presentations, your work ethic… you thought me so many things by example! Candice, you were such a ray of sunshine, with your big smile and free French lessons, besides having helped me in my experiments so many times! You will always be my Western blot queen! Audrey, the expression that always comes to my mind when I think of you is “we need more controls!!!”. You showed me a whole new level of experimental organization and planning, and that was absurdly helpful for me, especially when my PhD dove into those huge and complex experiments that required a lot of organization in order to work. Thank you for that. Eve, my dear Master student, having had the opportunity to supervise you was truly amazing and a unique experience! Our work together unleashed my biggest PhD project, and you became someone that I deeply care about and I am sure will rock this Science world so much! Nico, you had a joke always ready and you spent most of your time teasing me, but you were also one of the people I discussed more Science with, and that was really amazing! Thank you for all your advice, tips, words of encouragement and for your true and beautiful excitement about Science, specially your beloved T6SS. Also, thank you for all the music suggestions, it was great to have someone around that is as excited for music as I am! Sandrine (Isaac), I will always remember you as someone who was not afraid to ask questions is a big seminar even having just started a post-doc. I admire that so very much and I wish I could learn to be more like you! Your sharp mind, attention to detail and incredible dedication will always inspire me. Jorge, it was great to have another Portuguese-speaking buddy in the lab, even though I had to go through an adaptation phase to understand your accent. But of course, much more than that, you thought me the value of deep thinking, of organizing my thoughts and sometimes sharing them out loud with others until they made sense. I also thank you for all the time we spent outside of the lab, for our shared excitement for food and art and for being my friend. To all

the other amazing people that were in the lab with me for not so long, but whom I will always take with me in my heart: Tiziana, Julien, Anne-Catherine, Ivan, Thibault. Also, to the new people who recently started or to those who will soon come back in to lab: Simon, Mischa, Alexis. It will be awesome to have you around for the rest of the time I will spend in the lab!

Now, a special and bigger-than-the-world thank you to my two lab siblings: Nina and Leonardo. You became much more than PhD colleagues and lab mates to me. We constructed a real friendship based on trust, real connection and so much laughter. Thank you for always being there with a joke ready to be cracked, with a warm hug or chocolate to help in those slightly more difficult days, for giving me your attention when I needed, for being interested in my projects and in my life. Thank you for all the wonderful time we spent in the lab but also outside, by the lake, in trips, in parks, having a barbecue… You were my family away from home. I am so very lucky to have had you two by my side over these years, and I want to keep you in my life until we are grumpy old buddies sharing a laugh no matter where our lives take us. Nina, you started your PhD 9 months after me and what a gift you were! We shared so much of our lives, we helped each other, we laughed and travelled together… you are a true friend. We are so similar in so many ways and so different in others, that I think our styles matched perfectly to go through our PhDs with style. Also, you were always so deeply and truly interested in my experiments! Your curiosity and love for Microbiology were truly inspiring and helped motivating me to keep going. And by the way, who knew “your” Clément would get along so well with “my” Matheus and that in the end we would become a little team! Leozinho, you came later but you did it with a “bang”, of course, with all the associated drama and charisma. You showed me a new way of functioning as a human and scientist, which included asking for help more often, collaborating, being less shy, saying my opinion… I’m still a baby in learning these new ways, but you did teach me so much! Thank you for being so funny and so dramatic, with you I always felt I was living

inside one of those Brazilian soap operas. Thank you for your support, for believing in me and motivating me every step of the way!

I would also like to thank all of my friends in Brazil who were far away but so present in my life, especially Bruna, Felipe, Lilian, Josi, Lisiele, Paola, Laura, Larissa and Isadora. Thank you for cheering me up, for being so excited in every step of my PhD, for having the patience to keep our friendships even stronger despite the distance!

Now, I would like to thank my whole (though very tiny) family. My parents, Walter and Viviane, who are true superheroes. Our lives were far away from easy, but you showed me the power of resilience, of giving our best, of being ethical even if that meant being left behind, of being true to my principles. I would never get here without you. Thank you for being so excited for my career and for supporting me no matter what! My sister, Julia. My other half in this world, my partner in the lows and highs of our lives, my huge supporter. Thank you for always being there, for not letting me down, for lifting me up when things were a bit hard, for cheering more than anything when things worked out (even if you had no idea what were those crazy experiments I was talking about!). You were here with me all the way and I feel blessed to share the world with you. You are a rock star and I am so very proud of the women we became! My grandparents, who are so technological that we can communicate and share details about the weather and the current drought without any problems. Thank you for supporting me even though you don't really understand why I had to come so far away to get a job… My dear aunt Inge, who sent me messages every day, asking how I was, keeping track of whether I was working in the lab or from home, if I was doing experiments or writing my thesis… your care was priceless. My huge family-by-association, especially Sandra, Pedro, Soraia and Rodrigo. Thank you for all your love and excitement for my career, for trusting that I would take good care of your son/ brother so far away from home, and for making me feel part of the family.

A very important thank you goes to my cat, Carotte. He came into our lives less than a year into my PhD and he just made everything so much more fun and colorful, full of stories and fluffiness. Thanks, buddy, for making me company and for warming up my computer to make sure it would work properly.

My final and most deep thank you goes to my love, my husband Matheus. You boarded in this journey through the Atlantic without anything to gain for yourself. You came because you trusted me, you believed when I said I had to do this and, of course, because you love me. Little did we know what this experience would bring us. Being away from everything we know was a huge challenge in every single way, but we survived, together! This PhD was a huge milestone for my professional career but I dare saying that it was a turning point for our relationship as well. We faced some of our toughest moments while living here, but also our biggest joys! All of this made us stronger as a couple and as people too. You were my rock, my safe harbor, my biggest supporter and believer. As expected, you made this experience fruitful and important for you as well. You learned a new language, you made new friends, you had your own art exhibition, you gained clients who loved your work and creativity, you started hiking and running, you fell in love with nature, you learned how to do taxes declaration in a foreign country, how to play the piano and the bass, you created a music album… you are the most creative and dedicated person I know, and I couldn't ask for a better partner to live this experience with. And now we have a little part of the two of us growing in my "oven", a little bump that we are connecting with and who will teach us a whole new aspect of the human experience: parenthood. I cannot think of a better to way to end the largest chapter of my life than like this. I love you.

CURRICULUM VITAE

Personal information

Name	Natália Carolina Drebes Dörr
Date of birth	26.12.1987
Citizenship	Brazilian
Email	natalia.dorr@gmail.com
Address	Chemin de la Cigale, 2C - 1010 Lausanne, Switzerland

Education

Since July 2016

PhD thesis in Molecular Life Sciences

École Polytechnique Fédérale de Lausanne (EPFL)

The molecular treasure box of environmental *Vibrio cholerae* strains

Supervision: Prof. Melanie Blokesch

2013 – 2015

Master of Science in Genetics and Molecular Biology

Federal University of Rio Grande do Sul (UFRGS), Brazil

Wolbachia dynamics in natural populations of hosts from the *willistoni* subgroup of *Drosophila*

Supervision: Prof. Vera Lucia da Silva Valente Gaiesky
Prof. Victor Hugo Valiati

2006 - 2012

Bachelor of Science in Biology

University of Rio dos Sinos Valley (UNISINOS), Brazil

Work and research experience

Since July 2016 PhD student at the Global Health Institute (GHI), EPFL, Switzerland

2015 Technical Sales Consultant – Molecular Biology
Forlab Comércio de Produtos e Equipamentos para Laboratórios, Brazil

2013 – 2015 Master student at the Genetics department (CNPq fellowship), UFRGS, Brazil

2012 – 2013 Environmental issues technical assistant
Projeconsult Engeneering Ltda, Brazil

2011 – 2012	Undergraduate research intern (with fellowship) Molecular biology laboratory, UNISINOS, Brazil Role of *Wolbachia* in the evolutionary history of Neotropical *Drosophila* species Supervision: prof. Victor Hugo Valiati
2011	Vienna BioCenter Summer School intern (June – August) (with fellowship) Max F. Perutz Laboratories, Vienna BioCenter, Austria miRNA regulation Supervision: Dr. Silke Dorner (RNA Biology Department)
2010-2011	Biology intern Biology Expert Advice from the Public Prosecutor's Office, Brazil
2010	EPFL Summer Research Program intern (June – August) (with fellowship) Swiss Institute for Experimental Cancer Research (ISREC), EPFL, Switzerland The role of miRNA in Bicaudal-C-mediated translational silencing Supervision: Prof. Daniel Constam
2010	Cell and Molecular Biology Summer Research Program intern (January) Medicine Faculty, University of São Paulo, Brazil
2007 – 2011	Volunteer intern and Undergraduate research Intern (with fellowship) Microbiology laboratory, UNISINOS, Brazil Interaction of plant and microbial proteins applied to Rice Integrated Pest Management Supervision: Prof. Lidia Mariana Fiuza

Languages

Portuguese	Native
English	Fluent
French	B1

Skills

Scientific

- Sterile technique with biosafety level 2 organisms
- Molecular cloning
- Genetic engineering of *Vibrio cholerae*
- Protein harvesting and detection by SDS-PAGE, silver staining, western blot analysis
- DNA isolation
- DNA pull down
- RNA isolation, cDNA synthesis, qRT-PCR
- 5' RACE (Rapid Amplification of cDNA Ends)
- Sterile technique for cell culture (amoebae)

- Infection of mini-guts
- Fluorescence microscopy
- Confocal (time-lapse) microscopy
- *Drosophila* collection in nature
- Insect rearing: *Drosophila*, Lepidoptera (*Spodoptera frugiperda*, *Diatraea saccharalis*)
- *Drosophila* fitness assays: survival, fertility
- Bioassays of *Bacillus thuringiensis* and Bt Cry proteins with insects (*S. frugiperda*, *D. saccharalis*, *Oryzophagus oryzae*, *Tibraca limbativentris*, *Oebalus poecilus)*
- Planning, implementation and evaluation of experiments
- Communication of scientific results to expert/ non-expert audiences

Teaching

- Supervision of bachelor students during MSc
- Teaching assistance during MSc (Evolutionary Biology)
- Teaching assistance during PhD (Infection Biology, Molecular Biology practical course)
- Supervision and co-supervision of master students during PhD

IT

- Image analysis (ImageJ and Fiji, basics)
- Genome analysis (Geneious, basics)
- Microsoft Office (Word, Excel, PowerPoint)
- Adobe Illustrator
- Data representation and statistics (GraphPrism, basics)

Science Communication

- Atelier scientifique (EPFL)
- Journée gymnasien (EPFL)
- Scientific blogging
- Scientific podcasting
- Organization of scientific events:
  - Biology Students Meeting (RABU) (UNISINOS)
  - A Hitchhiker's Guide to the Biosphere (Introduction for a Bachelors in Biology) (UNISINOS)

Conferences and presentations (during PhD)

- Talk given to students from University of North Carolina in Asheville (March 2021; remotely) – oral presentation
- T6SS SympoZoom (April 2020; remotely) – oral presentation
- Gordon Research Conference and Seminar: Animal-Microbe Symbioses as Nested Ecosystems (June 2019, Mount Snow, USA) – poster presentation. With a LS2 (Life Sciences Switzerland) travel grant.
- SME19 (Swiss Microbial Ecology Meeting, January 2019, Lausanne) – oral presentation
- SSM18 (Swiss Society for Microbiology Meeting, August 2018, Lausanne) – oral presentation
- ISME17 (17th International Symposium on Microbial Ecology, August 2018, Leipzig) – poster presentation
- Symbioses Summer School (Instituto Gulbenkian de Ciência, July 2017, Oeiras, Portugal) – poster presentation

Publications

- **Drebes Dörr NC** and Blokesch M. Interbacterial competition and anti-predatory behavior of environmental *Vibrio cholerae* strains. ***Environ. Microbiol.*** 2020 22(10): 4485-4504. DOI: 10.1111/1462-2920.15224

- Van der Henst C, Vanhove AS, **Drebes Dörr NC**, Stutzmann S, Stoudmann C, Clerc S, Scrignari T, Maclachlan C, Knott G and Blokesch M. Molecular insights into *Vibrio cholerae*'s intra-amoebal host pathogen interactions. ***Nat Commun.*** 2018 9: 3460. DOI: 10.1038/s41467-018-05976-x.

- **Drebes Dörr NC** and Blokesch M. Bacterial type VI secretion system facilitates niche domination. Commentary. ***PNAS*** 2018 115 (36) 8855-8857. DOI: 10.1073/pnas.1812776115.

- Matthey N*, **Drebes Dörr NC*** and Blokesch M. Long-read-based genome sequences of pandemic and environmental *Vibrio cholerae* strains. ***Microbiol. Resour. Announc.*** 2018 7 (23) e01574-18. DOI: 10.1128/MRA.01574-18. (*equal contribution)

- Müller MJ, **Dörr NCD,** Deprá M, Schmitz HJ, Valiati VH, Valente VLS. Reevaluating the infection status by the *Wolbachia* endosymbiont in *Drosophila* neotropical species from the *willistoni* subgroup. ***Infect. Genet. Evol.*** 2013 19: 232-239. DOI: 10.1016/j.meegid.2013.07.022.

- Pinto LMN, **Dörr NC**, Ribeiro APA, Salles SM, Oliveira JV, Menezes VG, Fiuza LM. *Bacillus thuringiensis* monogenic strains: screening and interactions with insecticides used against rice pests. **Braz. J. Microbiol.** 2012 43 (2): 618-626. DOI: https://doi.org/10.1590/S1517-83822012000200025.

Interests

- Books
- Movies
- Music
- Art museums
- Wine, cooking, baking
- Fitness
- Travelling
- Knitting
- Gardening