Type VI Secretion System-Mediated Microbial Competition of *Vibrio cholerae*

by

Daniel Unterweger

A thesis submitted in partial fulfillment of the requirements for the degree of

Doctor of Philosophy

in

Bacteriology

Department of Medical Microbiology and Immunology
University of Alberta

Abstract

Vibrio cholerae is a Gram-negative bacterial species that consists of over 200 serogroups with differing pathogenic potential. Only O1 serogroup strains have been associated with the pandemic spread of cholera, a disease characterized by watery diarrhoea that can lead to severe dehydration and death. The current 7th cholera pandemic is still ongoing and follows six pandemics recorded during the 18th and 19th centuries. O1 serogroup strains make use of the main virulence factors – toxin coregulated pilus and cholera toxin – to colonize the small intestine of the human host and induce watery diarrhoea. Although strains not belonging to the O1 or O139 serogroups lack or do not effectively make use of these virulence factors, strains like the O37 serogroup strain V52 or the O39 serogroup strain AM-19226 can cause local outbreaks and cholera-like symptoms characterized by a mild form of diarrhoea. All *V. cholerae* strains sequenced to date harbour genes for the type VI secretion system (T6SS), a molecular puncturing device with homology to the tail-spike complex of the T4 bacteriophage that allows translocation of effectors into neighbouring eukaryotic and prokaryotic cells. Translocated effectors are toxic unless the target cell produces an immunity protein against the cognate effector. Proteins of the T6SS are encoded in three gene clusters on the large and small chromosome of *V. cholerae*. The T6SS of pandemic strains is activated in vivo during infection. *V. cholerae* employs its T6SS to kill other prokaryotic cells using effectors with peptidoglycan-degrading, pore-forming, and lipase activities. The T6SS also confers virulence toward eukaryotic phagocytic cells such as amoebae and macrophages. An effector with an actin-crosslinking domain is toxic to eukaryotic cells.

The T6SS is differentially regulated among *V. cholerae* strains. The O37 serogroup strain V52 has an active T6SS under laboratory conditions. The O1 serogroup strain C6706 represses its T6SS under laboratory conditions and depends on inducing signals like mucin to activate the secretion system. Immunity protein-encoding genes are differentially regulated than the remaining genes of the T6SS gene clusters and are expressed under laboratory conditions in the O1 serogroup strain C6706. As a result, pandemic *V. cholerae* O1 serogroup strain C6706 is immune to a T6SS-mediated attack by the O37 serogroup strain V52. Three immunity proteins TsiV1, TsiV2 and TsiV3 protect from the cognate T6SS effectors TseL, VasX and VgrG-3, respectively.

Characterization of environmental *V. cholerae* strains isolated from the Rio Grande Delta revealed intraspecific competition of *V. cholerae* such that *V. cholerae* strains kill each other in a T6SS-dependent and -independent manner. Bioinformatics analyses of 37 *V. cholerae* strains indicated that effector and immunity protein-encoding genes encoded within otherwise conserved gene clusters differ widely among *V. cholerae* strains. Effector modules were defined that contain effector and immunity protein-encoding genes which are likely exchanged between *V. cholerae* strains. The set of effector modules determines the compatibility group of a *V. cholerae* strain. Strains of the same compatibility group are able to coexist in direct contact and do not kill each other in a T6SS-dependent manner, because immunity proteins protect from cognate effectors. Strains of different compatibility groups kill each other in a T6SS-dependent manner because of the inability of the immunity proteins to protect from the effectors. All analyzed pandemic strains can be assigned to a single compatibility group and are thus expected to coexist among each other but not with nonpandemic strains. Further analysis

of pandemic strains revealed that strains responsible for the sixth and likely also for preceding pandemics are weak microbial competitors.

The bioinformatics analysis of *V. cholerae* strains also revealed diverse T6SS effectors. A chimeric adaptor protein was identified and characterized that mediates the translocation of diverse effectors of the T6SS encoded in modules. After T6SS-mediated effector export is completed, Tap-1 is retained in the bacterial cell to load other T6SS machines.

The presented work advances our understanding of the diversity of T6SS effector and immunity proteins among *V. cholerae* strains, and proposes a model for the translocation of diverse effectors via adaptor proteins. The T6SS effector module sets could be targeted as a biomarker for pandemic strain identification for epidemic risk assessment. Classification of the *V. cholerae* strains based on effector module sets may affect future outbreak management, diagnostics, and therapeutic approaches.

Preface

Research conducted in this thesis forms part of a research collaboration with Daniele Provenzano (University of Texas, Brownsville, TX, USA). Daniele Provenzano provided the Rio Grande *V. cholerae* isolates that I tested for intraspecific competition (chapter 3). The screen of the Keio library (appendix B) was performed in collaboration with Nicole Acosta from Tracy Raivio's research group (Department of Biosciences, University of Alberta, AB, Canada). Nicole Acosta and I screened the Keio collection and contributed equally to this project. Tracy Raivio contributed to the supervision of this project.

Parts of chapter 2 of this thesis were published as Miyata, S.T., Unterweger, D., Rudko, S. and Pukatzki, S., Dual expression profile of type VI secretion system immunity genes protects pandemic *Vibrio cholerae*. *PLoS Pathogens*, 2013. 9(12): p. e1003752. I contributed to the design of experiments, performed experiments and interpreted data of this article. Sarah Miyata contributed to the design of experiments, performed experiments, interpreted data and wrote the draft of the manuscript. Sydney Rudko contributed to the design of experiments, performed experiments and interpreted data. Stefan Pukatzki was the supervisory author, designed and performed experiments, interpreted data and edited the manuscript.

Parts of chapter 3 were published as Unterweger, D.*, M. Kitaoka*, S.T. Miyata, V. Bachmann, T.M. Brooks, J. Moloney, O. Sosa, D. Silva, J. Duran-Gonzalez, D. Provenzano and S. Pukatzki, Constitutive type VI secretion system expression gives *Vibrio cholerae* intra- and interspecific competitive advantages. PLoS One, 2012. 7(10):e48320. Maya Kitaoka and I contributed equally to this manuscript. I designed and performed experiments, interpreted data and contributed to the draft of the manuscript. Stefan Pukatzki was the supervisory author, designed and performed experiments, interpreted data and edited the manuscript. The remaining authors designed and performed experiments and interpreted data.

Parts of chapter 4 were published as Unterweger, D., Miyata, S.T., Bachmann, V., Mullins, T., Brooks, T., Kostiuk, B., Provenzano D. and Pukatzki, S., The *Vibrio*

cholerae type VI secretion system employs diverse effector modules for intraspecific competition. *Nature Communications*, 2014. 5:3549. I designed and performed experiments, interpreted data and wrote the draft of the manuscript. Stefan Pukatzki was the supervisory author, designed and performed experiments, interpreted data and edited the manuscript. The remaining authors designed and performed experiments and interpreted data.

Parts of chapter 5 were published as Unterweger D., Kostiuk, B., Oetjengerdes, R., Wilton, A., Diaz-Satizabal, L., Pukatzki, S., Chimeric adaptor protein Tap-1 translocates diverse type VI secretion system effectors in *Vibrio cholerae*. EMBO J accepted. I designed and performed experiments, interpreted data and wrote the draft of the manuscript. Stefan Pukatzki was the supervisory author, designed and performed experiments, interpreted data and edited the manuscript. The remaining authors designed and performed experiments and interpreted data.

Animal studies used in this study (AUP0000320) were reviewed and approved by the Animal Care and Use Committee – Health Sciences at the University of Alberta. This committee adheres to the policies and standards of the Canadian Council on Animal Care. The University of Alberta's Animal Welfare Assurance Number is #A5070-01.

Acknowledgments

I would like to thank my supervisor, Stefan Pukatzki, for the opportunity to host me in his lab for the last five years and for a challenging, exciting and satisfying project. I enjoyed the freedom to pursue projects that I was interested in and to do so in a sometimes unconventional form. I thank Stefan for sharing his wisdom, being extremely generous, the open door, and for long discussions about new results and about how to overcome the next hurdle. I am grateful for Stefan's excitement for each of the projects and his dedication to publish the results despite demanding reviewers. To each of the publications, Stefan also contributed with experimental data. During the last years, I had outstanding opportunities to attend international conferences and to visit other labs. I am grateful to Stefan for his support in finding a place and in securing a scholarship to pursue a postdoctoral fellowship. I also would like to thank the remaining members of the Pukatzki family, Anna Manalo and Georgia, for welcoming me in Edmonton and supporting me throughout this project.

All of the here described projects are products of the entire lab. Maya Kitaoka, Sarah Miyata, Teresa Brooks, Verena Bachmann, Travis Mullin, Ben Kostiuk and Laura Satizabal-Diaz have been crucial for completing this work. I thank them for their help, discussions, patience, generosity and active participation in the projects.

Over the last years, I helped the summer students Miriam Tulatz, Jan Leipert, Rina Oetjengerdes and the MMI499 student Ashley Wilton. I thank them for their patience and dedication.

I would like to thank our collaborator and friend Daniele Provenzano for helpful discussions, motivation and for sharing the RGVC isolates with us. To test the concept of compatibility experimentally, we depended on additional strains that were gratefully provided by Jun (Jay) Zhu, Michelle Dziejman and Shelly Payne. I would like to thank everybody else who contributed to our publications in various forms.

Tracy Raivio and Rob Ingham joined my supervisory committee four years ago. Thank you for your input and mentorship. I would like to thank Eytan Wine and Andrew Camilli for acting as external examiners in my thesis defence.

I would like to thank Alberta Innovates - Health Solution and the German National Academic Foundation for fellowships. The Canadian Institutes of Health Research, the German Association for General and Applied Microbiology and the Faculty of Graduate Studies and Research supported me with travel awards. I would like to thank all other funding sources that contributed to the project and the completion of the degree.

I would like to thank the Department of Medical Microbiology and Immunology for admitting me to their PhD program.

Thank you to my family and friends for all their help and support with scientific and non-scientific matters.

I would like to thank Stefan Pukatzki and Marcia Craig for their comments and edits during the preparation of this thesis.

Table of Contents

List of Tables

List of Figures

Abbreviations

ACD	Actin cross-linking domain
BLAST	Basic local alignment search tool
CF	cystic fibrosis
CFU	Colony forming units
CT	cholera toxin
DNA	Deoxyribonucleic acid
GFP	Green fluorescent protein
HCD	High cell density
Hcp	Haemolysin co-regulated protein
HRP	Horseradish peroxidase
kDa	Kilodaltons
LB	Luria-Bertani
LCD	Low cell density
LPS	Lipopolysaccharide
OD	Optical density
PBD	Peptidoglycan-binding domain
PBS	Phosphate buffered saline
PCR	Polymerase chain reaction
RGVC	Rio Grande *Vibrio cholerae*
SDS	Sodium dodecyl sulfate
SDS-PAGE	Sodium dodecyl sulfate polyacrylamide gel electrophoresis
Sec	General secretion
T1SS	Type I secretion system
T2SS	Type II secretion system
T3SS	Type III secretion system
T4SS	Type IV secretion system
T5SS	Type V secretion system
T6SS	Type VI secretion system
T7SS	Type VII secretion system

TA	Toxin/anti-toxin
Tap	Type VI secretion system adaptor protein
TCP	Toxin co-regulated pilus
TsiV	Type VI secretion system immunity of *Vibrio cholerae*
Vas	Virulence-associated secretion
VgrG	Valine glycine repeat protein G
VPI	Vibrio pathogenicity island
X-Gal	5-bromo-4-chloro-3-indolyl-β-D-galactopyranoside

Chapter 1
Introduction

1 Introduction

1.1 The marine bacterium *Vibrio cholerae* and the disease cholera

The Gram-negative bacterium *V. cholerae* is diverse with over 200 serogroups. Two serogroup strains, O1 and O139, are associated with cholera, a disease characterized by watery diarrhoea and severe dehydration. This section summarizes the hallmarks of the disease cholera and introduces the pathogenic characteristics of *V. cholerae*.

1.1.1 *Vibrio cholerae* is abundant in the marine environment

Vibrio cholerae belongs to the genus of *Vibrionaceae*. Among the five classes within the phylum proteobacteria, *V. cholerae* belongs to the gammaproteobacteria. As such, it has the shape of a curved rod and exhibits a polar flagellum. *V. cholerae* has been isolated from marine environmental samples on multiple continents. Strain 623-39 from Asia was isolated in Bangladesh in 2002, a European strain VL426 was isolated in Germany, strain BX330286 was isolated in Australia in 1986, the North American strain 2740-80 was isolated in the USA in 1980, and the South American strain 1587 was isolated in Peru in 1994.

In the marine environment, *V. cholerae* is associated with zooplankton and phytoplankton on which it forms biofilms (Rawlings et al., 2007; Tamplin et al., 1990). In such biofilms, bacteria can use chitin (a derivative of glucose) as a single carbon source. As discussed in chapter 7.3 of this thesis, chitin also renders *V. cholerae* naturally competent and activates the type VI secretion system of *V. cholerae* (Borgeaud et al., 2015). In the environment, *V. cholerae* can enter a condition described as viable but nonculturable (VBNC) or active but nonculturable (ABNC) (Colwell, 2000; Kell et al., 1998; Xu et al., 1982). Bacteria in this state cannot be cultured despite being metabolically active (Colwell, 2000).

Both phenotype and the genotype diversity can be observed in *V. cholerae* strains. Phenotypically, *V. cholerae* strains can be divided into more than 200 serogroups based on their surface O-antigen. Whole genome analysis revealed differences between the genomes of individual strains (Chun et al., 2009). Chun and colleagues compared the genomes of 23 *V. cholerae* strains from 12 different serogroups; 2,432 genes were

present in all analyzed genomes. This number corresponds to 69.7% of the genes on the large chromosome and 59.9% of the genes on the small chromosome of the O1 serogroup strain N16961 (Chun et al., 2009). All remaining genes were absent or replaced by different genes in the genomes of other strains. Some of the genes that are not found in the genomes of all *V. cholerae* strains encode virulence factors. Only O1 serogroup strains have been associated with pandemic cholera.

1.1.2 The epidemiology of pandemic cholera

The identification and characterization of the bacterium *V. cholerae* is tightly connected to the disease it causes, which has killed millions of people worldwide since its global spread in the 19th century.

The cause of cholera was unknown until John Snow and Henry Whitehead conducted studies on the source of a cholera outbreak in London in 1854 (Newsom, 2006). Based on their observation on the spread of the disease and the association of cholera cases with contaminated water wells, Snow and Whitehead suggested that cholera is a water-borne disease transmitted by the fecal-oral route. The first to identify the bacterium *V. cholerae* in the stool of cholera patients was the Italian medical student Filippo Paccini in 1854 (Pacini, 1854). Robert Koch then isolated *V. cholerae* from patients in Egypt in 1883 and in India in 1884 (Koch, 1894). Recent advances in the ability to sequence the whole genome of *V. cholerae* strains have facilitated the analysis of the diversity within the species. Sequencing of ancient DNA sequences from conserved intestines of cholera patients has enabled us to determine the genome of *V. cholerae* strains we cannot culture any more. A recent example is the genome sequence of the 2nd pandemic strain PA1849 from an outbreak in Philadelphia in 1849 (Devault et al., 2014).

Most of what we know about pandemic *V. cholerae* strains that belong to the O1 or O139 serogroup is derived from strains of the sixth and seventh pandemic. The first of six pandemics, which all originated on the Indian subcontinent, lasted from 1817 to 1823 (Kamal, 1974; Pollitzer, 1959a). In previous years cholera had been endemic in the lower Ganges Delta, where pilgrim travel, an increasing shipping trade, and army movements

likely contributed to the first pandemic throughout India, eventually spreading the disease to China, Japan, and the Persian Gulf (Hays, 2005). The second pandemic, 1829–1851, was the first pandemic to reach North America. Immigrants travelling on ships from Ireland to Montreal, Canada, brought cholera, and the disease spread south to New York, Philadelphia, Baltimore, and Washington (Chambers, 1938a, b). The third pandemic, 1852–1859, was studied by John Snow and Henry Whitehead in England. In North America, settlers moved west, and so did cholera. At the beginning of the fourth pandemic, 1863–1879, cholera reached New Orleans and cities along the Mississippi and Ohio rivers (Billings J.S., 1975). During the fifth pandemic, 1881–1896, Robert Koch published his observations of the motile cholera bacterium (Koch, 1894). This pandemic also affected the South American countries Argentina, Chile, and Peru (Laval, 1989). The sixth pandemic occurred between 1899 and 1923 and, like the previous cholera pandemics, was probably caused by strains of the classical biotype (Barua, 1992). This pandemic heavily affected populations in the Middle East and on the Balkan peninsula (Pollitzer, 1959a). The seventh pandemic arose in Indonesia in 1961 and is caused by El Tor biotype strains. The name of the biotype is derived from the city El Tor in Egypt, where the first isolate of this biotype was described in 1905. O1 serogroup strains of the classical and El Tor biotypes are subdivided into three serotypes—Ogawa, Inaba, and Hikojima—based on the surface antigens. The Hikojima seroypte is very rare (Sakazaki and Tamura, 1971). Strains of the Ogawa serotype express the surface antigens A and B, strains of the Inaba serotype possess the antigens A and C, and strains of the Hikojima serotype possess antigens A, B, and C (Burrows et al., 1946). The B antigen contains a methyl group on the terminal sugar of the O-antigen, the C antigen does not (Villeneuve et al., 2000). A combination of core and O-antigen polysaccharides forms the A antigen (Villeneuve et al., 1999). *V. cholerae* of the O139 serogroup emerged in the Bay of Bengal in India around 1992–1993 (Faruque et al., 2003b). These virulent strains likely emerged through horizontal gene transfer between O1 and O139 serogroup strains (Bik et al., 1995). Because of the rapid spread of O139 strains into Thailand, Malaysia, Pakistan, and Nepal, O139 serogroup strains were temporarily called the 8th pandemic (Swerdlow and Ries, 1993). However, this outbreak remained endemic.

Strains of the classical and El Tor biotypes differ in their phenotype, genotype, and virulence potential. Tests have been developed to differentiate classical from El Tor strains based on their sensitivity to lytic phages and their ability to haemolyse erythrocytes. Classical strains are lysed by classical type IV bacteriophage. El Tor strains haemolyse sheep erythrocytes (Kaper et al., 1995). Bioinformatics analyses suggest that strains of the El Tor and classical biotype are derived from the same common ancestor, although their genomes differ. Classical biotype strains cause more severe disease symptoms than El Tor strains in humans, likely because the former produce higher levels of cholera toxin (Gangarosa, 1974). The reason for this might be in the different alleles of *ctxB* they encode or in the different regulation of their virulence factors (Jonson, 1990; Kaper et al., 1995). Recently, hybrid strains emerged that are similar to strains of the El Tor biotype but encode the *ctx* allele of classical strains (Safa et al., 2010). Because the classical *ctx* allele causes more sever disease symptoms, these hybrid El Tor strains are more virulent than their wild-type El Tor counterparts.

How classical strains got replaced by El Tor strains in causing pandemic cholera remains to be determined. Classical biotype strains are still sporadically isolated from the environment. The El Tor biotype strains originate from the Bay of Bengal and disseminated in the second half of the 20th century world-wide in at least three independent but overlapping waves (Mutreja et al., 2011). Mutreja and colleagues performed extensive phylogenetics analyses based on whole genome sequences of 154 *V. cholerae* strains to further characterize the origin and dissemination of the 7th pandemic (Mutreja et al., 2011). Only the first of the three waves spread to South America after passing Africa between 1981 and 1985 (Mutreja et al., 2011). Strains that spread from Africa to South America differ from all other analyzed strains by the acquisition of the genomic islands VSP-2 and WASA1 (Mutreja et al., 2011; O'Shea et al., 2004). The functions of the proteins encoded in genes on these islands need to be further elucidated. In addition to Africa, *V. cholerae* of this first wave spread throughout Asia and Europe. *V. cholerae* strains of the second (1990–2002) and third (1989 to the present) waves differ from the strains of the first wave by the acquisition of the integrative and conjugative element SXT/R391 that carries antibiotic resistance (Carraro and Burrus, 2014; Mutreja et al., 2011). Waves 2 and 3 are more geographically restricted than the

first wave; with the exception of Haiti, epidemics have been restricted to Africa and South Asia since 2003.

In 2010, a cholera outbreak in Haiti resulted in close to 700,000 cholera cases and over 8,000 deaths after the introduction of pandemic cholera by a soldier from Nepal who was dispatched to the island in response to an earthquake (Orata et al., 2014). *V. cholerae* is now endemic in Haiti. The most recent outbreak reported by the World Health Organization (WHO) was in South Sudan early in 2014 (WHO, 2014a).

1.1.3 The disease cholera is characterized by watery diarrhoea

Cholera patients acquire *V. cholerae* through the uptake of contaminated water or food and show symptoms of watery diarrhoea of up to 1 litre per hour which results in severe dehydration and can lead to death. When left untreated, the mortality rate of cholera is 50% (WHO, 2010). Standard treatment is rehydration therapy. The World Health Organization (WHO) recommended a solution of sodium, potassium, chloride, citrate, and glucose to rehydrate patients and to accommodate for electrolyte loss (WHO, 2006). These oral rehydration solutions successfully treat up to 80% of patients (WHO, 2008). In severe cases, antibiotics can be administered. This treatment option is limited by an increasing number of antibiotic-resistant strains isolated over the last years (Faruque et al., 2006; Kim et al., 2010; Kitaoka et al., 2011b).

Cholera vaccines are one approach to protect people at risk. The human immune system has the potential to protect from cholera, as indicated by a study on healthy volunteers that developed a three year protective immune response (Cash et al., 1974a; Levine et al., 1981; Levine et al., 1979). Lipopolysaccharide (LPS) of *V. cholerae* induces an innate and humoral immune response in infected individuals (Finkelstein, 1962; Flach et al., 2007; Neoh and Rowley, 1970). The response of the adaptive immune system is primarily mediated by B-cells and to a small extent by T-cells. This imbalanced response might be one reason for the relatively short duration of three years of the protective immune response (Harris et al., 2009). Because the LPS O-antigen varies between *V. cholerae* serotypes Inaba, Ogawa, and Hikojima (see section 1.2 of this chapter for details), immunity to one serotype was found to be only partially protective to

other serotypes (Cash et al., 1974a; McCormack et al., 1969). Also, patients with different blood groups differed in the strength of the immune response, those with the blood group 0 having the weakest immune response (Clemens et al., 1989; Harris et al., 2005; Provenzano et al., 2006). During the last 50 years, cholera vaccines containing live-attenuated bacteria and whole-cell killed bacteria were developed (Bishop and Camilli, 2011). A vaccine containing whole-cell killed bacteria administered orally in a volunteer study induced a three-year long protection, similar to an infection by wild-type bacteria with an efficacy of 50% (Clemens et al., 1990). A live attenuated vaccine in which the pathogen was live but its virulence was attenuated, was developed based on the classical O1 Inaba strain O569B. Most of the *ctxA* was deleted in this strain. A comparison of the immune response of vaccinated individuals to the immune response of previously infected individuals revealed that the vaccination was less robust than the natural infection (Levine et al., 1988a). A major challenge for future developments is a vaccine that induces a protective immune response in children (Bishop and Camilli, 2011).

The best prevention of cholera spread remains safe drinking water and adequate sanitation. In areas without access to clean water, inexpensive methods such as water filtration to eliminate plankton-associated *V. cholerae* bacteria have been shown to be effective and to reduce the number of bacteria by 99% (Huq et al., 1996). Despite these straightforward measures to contain *V. cholerae*, up to 5 million cholera cases and up to 120,000 deaths due to cholera around the globe annually are estimated (WHO, 2014c).

1.1.4 Virulence factors of *V. cholerae* and the pathophysiology of cholera

Once *V. cholerae* is taken up orally, many bacteria succumb to the low pH in the stomach. A study in which increasing doses of *V. cholerae* were administered to prison inmates showed that the infectious dose dropped from 10^{11} to 10^{6} bacteria when the stomach acid was neutralized with sodium bicarbonate (Cash et al., 1974b). Merrell and colleagues identified genes required for the acid tolerance response of *V. cholerae* (Merrell et al., 2002b). Mutants with transposon insertions in these genes were identified as part of a signature tagged mutagenesis screen for colonization–deficient mutants in the

infant mouse. In mice infected with mutants of *ghsB*, *hepA*, or *recO*, which are implicated in the acid tolerance response, 1000 fold fewer bacteria were recovered after a 24 hour infection compared to infection with the wild-type strain N16961 (Merrell et al., 2002b). *V. cholerae* exposed to acid prior to infection of the infant mouse outcompeted untreated bacteria (Merrell and Camilli, 1999).

On reaching the small intestine, *V. cholerae* makes use of its two main virulence factors, toxin coregulated pilus (Tcp) and cholera toxin (CT), to colonize the small intestine and induce water efflux. Taylor and colleagues identified a transposon *V. cholerae* mutant with dysfunctional *tcp* that colonized the infant mouse at reduced rates (Taylor et al., 1987). The toxin coregulated pilus (Tcp) belongs to the family of type IV pili and requires at least 15 proteins for its assembly and function (Iredell and Manning, 1994). The locus that contains the *tcp* genes is referred to as *Vibrio* pathogenicity island 1. This island differs from the remaining chromosome in its GC content, contains integrase- and transposase-encoding genes, and is flanked by *att* sites (Karaolis et al., 1999). These are indications of its acquisition by horizontal gene transfer, possibly by a bacteriophage (Karaolis et al., 1999), but questioned by Faruque and colleagues (Faruque et al., 2003c). Tcp might facilitate the attachment of *V. cholerae* bacteria to each other to form microcolonies (Kirn et al., 2000) and to the epithelium (Krebs and Taylor, 2011). In addition, Tcp acts as a receptor for the CTX phage that provides pandemic strains with *ctxAB*, the genes encoding cholera toxin (Waldor and Mekalanos, 1996). Subsequently, phage transduction requires conditions under which *V. cholerae* expresses Tcp, such as the in vivo environment of the intestine during infection. In infant mice intragastrically infected with a donor strain, which contains a kanamycin-marked CTX phage and the 7th pandemic strain Bah-2, 0.5% of Bah-2 that was recovered from the intestine after 24 hours was transduced (Waldor and Mekalanos, 1996). Phages like the CTX phage can contribute to the virulence of *V. cholerae* by equipping the bacterium with virulence factors. However, phages can also interfere with the virulence of *V. cholerae* as shown for the lytic phage ICP2 which selects in vivo for resistant *V. cholerae* mutants (Seed et al., 2014). Kim Seed and colleagues observed mutations in the outer membrane protein OmpU and ToxR, a regulator of OmpU, and virulence factors (as described in detail later in this section), among *V. cholerae* isolates resistant to the phage ICP2. In rabbits

inoculated with phage and equal amounts of phage-sensitive wild-type *V. cholerae* and phage-resistant mutant, the phage-resistant mutant was recovered at 10,000 fold higher numbers than the wild-type strain (Seed et al., 2014), indicating that the phage selects for resistant bacteria in vivo. In an animal experiment, phage-resistant *toxR* mutants colonized the small intestine of an infant mouse at 100 to 1000 fold fewer numbers than wild-type bacteria (Seed et al., 2014). This result indicates that resistance to phage might come at a cost of attenuated virulence.

Cholera toxin consists of two subunits, CtxA and CtxB (Sixma et al., 1992). On release of the two proteins by the type II secretion system (Sandkvist et al., 2000), CtxB interacts with GM1 gangliosides on eukaryotic cells and mediates the uptake of CtxA into the cell (King and Van Heyningen, 1973). CtxA transfers an ADP-ribose moiety and NAD to the alpha-subunit of the G_s protein (Gill and King, 1975). The thereby ADP-ribosylated alpha subunit subsequently activates adenylate cyclase (Cassel and Selinger, 1977; Kahn and Gilman, 1984). Increased levels of cAMP result in less sodium uptake and increased chloride efflux through the CFTR channel, which is activated by protein kinase A (Cassel and Pfeuffer, 1978; Gill and Meren, 1978). This imbalance of ions results in an osmotic gradient that promotes water efflux into the intestinal lumen and results in watery diarrhoea, the characteristic symptom of cholera.

Tcp and Ctx are regulated by the ToxR regulon. ToxR is a transmembrane protein and directly controls expression of *ctxAB* (Miller et al., 1987). The activity of ToxR on the promoter of *ctxAB* was identified in a screen based on an *E. coli* strain with a transcriptional fusion of the promoter of *ctxAB* to lacZ. A genomic library of *V. cholerae* was transformed into this *E. coli* strain, which turned blue in the presence of *toxR* when grown on X-Gal (Miller and Mekalanos, 1984). ToxS is required to stabilize ToxR, which dimerizes upon activation by external stimuli (DiRita and Mekalanos, 1991). ToxR also induces expression of ToxT, another member of the ToxR regulon and an AraC-like transcriptional activator (Higgins et al., 1992). ToxT also induces expression of *ctxAB* and additionally activates Tcp. The gene *toxT* is encoded downstream of the genes that encode proteins of the toxin coregulated pilus (Higgins et al., 1992). A fine-tuned response to the intestinal microenvironment likely enables *V. cholerae* to produce Tcp early during infection and later cholera toxin and Tcp in close proximity to the epithelium

in the intestinal crypts (Lee et al., 1999; Nielsen et al., 2010). Bile is present at high concentrations in the intestinal lumen and decreases expression of *tcp* and *ctxAB* (Gupta and Chowdhury, 1997; Schuhmacher and Klose, 1999). Bicarbonate is present at increasing concentrations close to the epithelium and induces the expression of *tcp* and *ctxAB* (Abuaita and Withey, 2009; Dietz and Field, 1973). In the laboratory, classical biotype *V. cholerae* strains express Tcp and CT when grown in LB at pH 6.5, 30 ºC; the El Tor biotype strains are grown in AKI media in standing and afterwards shaking culture (Gardel and Mekalanos, 1996; Iwanaga et al., 1986).

The role of TCP and CTX as main virulence factors is supported by the results of a study on human volunteers, which reports a colonization defect of a *V. cholerae* strain deficient in *tcpA* and attenuated disease symptoms of a strain deficient in *ctxA* (Herrington et al., 1988). *V. cholerae* gets flushed out of the body in the watery diarrhoea of the infected individual, and as it permeates the environment it is transmitted to new hosts via the fecal-oral route. In the environment, genes that were important during infection become less important and genes important for fitness in the environment are activated (Kamp et al., 2013).

When exiting the human host, *V. cholerae* is in a transient and reversible hyperinfectious state. *V. cholerae* from the stool of cholera patients outcompeted *V. cholerae* grown in vitro or in pond water by up to 100 fold (Merrell et al., 2002a). The hyperinfectious state is characterized by high levels of motility and nutrient acquisition. Pond water enhances the infectivity of *V. cholerae*, as shown by the following experiment of Nelson and colleagues. *V. cholerae* incubated in pond water for 5 hours had a 20 fold lower infectious dose than *V. cholerae* not incubated in pond water in intragastrically infected infant mice (Nelson et al., 2008). In the environment, *V. cholerae* can enter a viable but not culturable state but remains infective. A study on human volunteers suggests that viable but nonculturable *V. cholerae* are capable of colonizing the intestine and inducing diarrhoea (Colwell et al., 1996).

V. cholerae strains in which *ctx* was deleted for experimental purposes or non-O1/non-O139 strains that naturally do not contain CTX still induce watery diarrhoea, but at 20 times lower volumes than cholera toxin positive strains (Levine et al., 1988b; Pollitzer, 1959b). This effect has been attributed to additional toxins of *V. cholerae* such

as HlyA, RtxA, and HapA that also contribute to prolonged colonization in the infant mouse (Olivier et al., 2007). HlyA is a cytotoxic haemolysin that causes fluid accumulation in the ligated rabbit illeal loop (Ichinose et al., 1987). RtxA contains multiple domains with antieukaryotic activity (Prochazkova et al., 2009). One of them is an actin-crosslinking domain (ACD) that mediates cell rounding of epithelial cells (Sheahan et al., 2004). HapA is a zinc-metalloprotease that disturbs the integrity of a eukaryotic cell layer and causes bloody diarrhoea (Hase and Finkelstein, 1991; Wu et al., 1996). Thus, *V. cholerae* employs a variety of toxins which are detrimental to the host during infection.

Another factor that contributes to *V. cholerae* virulence is motility (Guentzel and Berry, 1975). Mutants deficient in the flagellum or the flagellar motor protein were outcompeted ~ 10 fold in the infant mouse by wild-type strains (Lee et al., 2001). Mutants able to rotate the flagellum only counter-clockwise and thus swim straight outcompeted mutants able to rotate the flagellum only counterclockwise and thus swim in zig-zag directions (Butler and Camilli, 2004). The role of motility has been implicated in *V. cholerae* penetration of the mucus layer, contact with the epithelium, and putatively in detachment during shedding upon multiplication (Mekalanos, 1995). Recent findings by Millet and colleagues indicated that motility is required for colonization of the proximal and medial part but not the distal part of the mouse intestine (Millet et al., 2014). *V. cholerae* bacteria in rice water stool (a watery stool containing white flecks of mucus, epithelial cells, and bacteria) are highly mobile, although nonchemotactic (Merrell et al., 2002a). Motility is one feature of the transient hyperinfectious state *V. cholerae* bacteria are in once they exit the host (Merrell et al., 2002a). Chemotaxis itself reduces infectivity (Fu et al., 2013) and is likely more important for *V. cholerae* in the environment (Kamp et al., 2013).

Individual virulence factors are tightly regulated and activated in a coordinated manner during infection. One trigger of virulence gene expression is cell density which is detected by the quorum sensing system. High concentrations of the autoinducers CAI-1 and AI-2 indicate high cell density and mediate the dephosphorylation of LuxO (Ng and Bassler, 2009). Subsequent repression of HapR results in a repression of virulence factor production (Zhu et al., 2002).

1.2 Bacterial secretion systems

At least four different secretion systems have been described in *V. cholerae*. One of them is the type II secretion system (T2SS) which is used to secrete cholera toxin into the extracellular space (Figure 1-1) (Reichow et al., 2010). The type III secretion system (T3SS) is found only in selected *V. cholerae* strains. One of the strains with a T3SS is the O39 serogroup strain AM-19226, which does not have *tcp* or *ctx* and requires its secretion system to colonize the small intestine of infant mice (Chaand et al., 2015; Dziejman et al., 2005; Shin et al., 2011; Tam et al., 2007). *V. cholerae* strains with Tcp can use this type IV pilus to secrete TcpF (Megli et al., 2011). Mutants of the *V. cholerae* strain O395 deficient in *tcpF* are recovered from the small intestine of infected infant mice at 1,000 fold lower numbers than wild-type O395 (Megli et al., 2011). The activity of TcpF remains to be determined. The gene cluster of the type VI secretion system is found in all sequenced *V. cholerae* strains to date (Zinnaka and Carpenter, 1972) and is the subject of this thesis. Its role in *V. cholerae* is discussed in detail in section 1.5. In this section I review common features of bacterial secretion systems found among Gram-negative bacteria and point out differences among them. I divide the secretion systems into two groups, those that span the inner and outer cell membrane, and those that span only the inner or outer cell membrane. Gram-positive bacteria only have one cell membrane but must cross an outer cell wall to transport proteins into the extracellular space (Abdallah et al., 2007). A type VII secretion system is used by various Gram-positive bacteria to transport proteins into the extracellular space. Among these bacterial species is *Mycobacterium tuberculosis* and the T7SS is associated with its virulence (Guinn et al., 2004; Stanley et al., 2003).

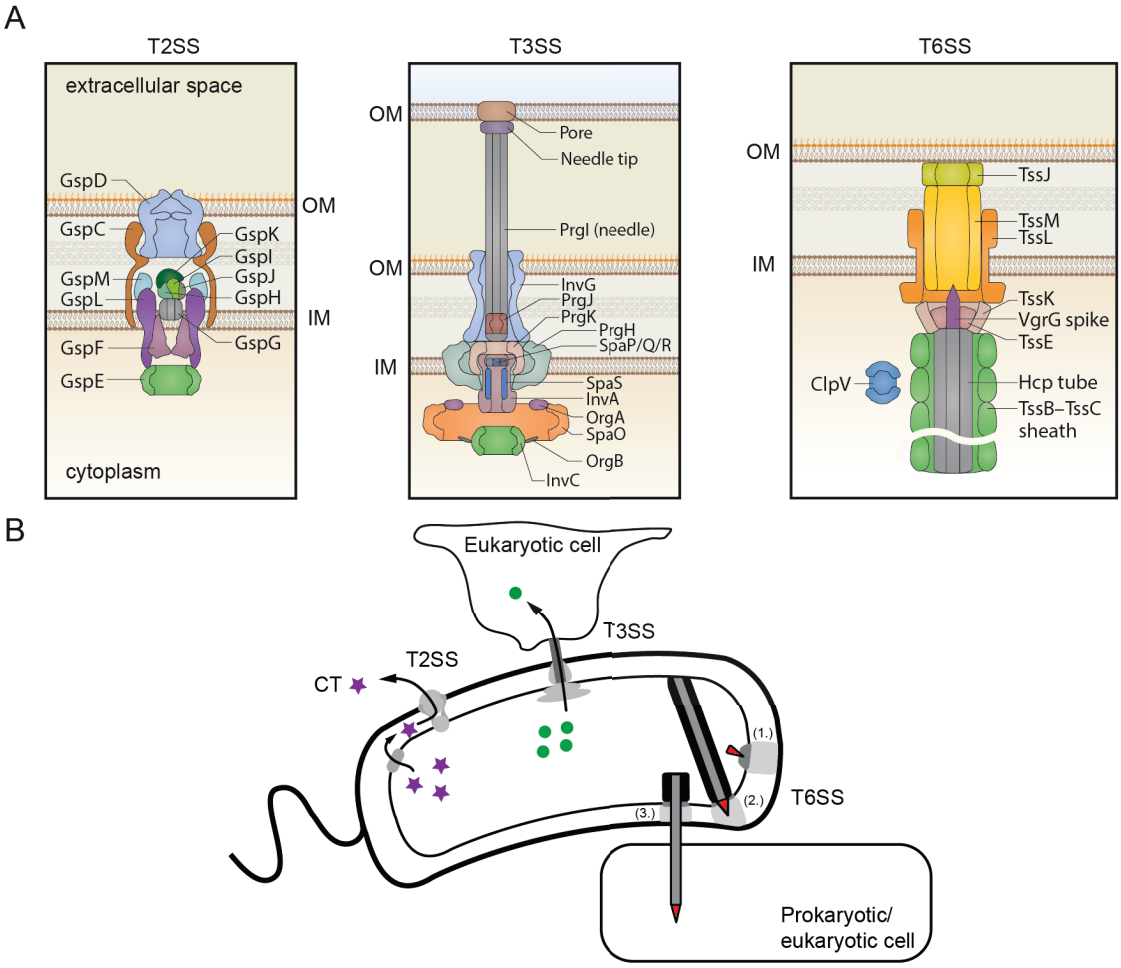


Figure 1-1 Type II, III, and VI secretion systems of *V. cholerae.*

(A) Graphical depiction of type II, III, and VI bacterial secretion systems. Reprinted by permission from Macmillan Publishers Ltd: Nature Reviews Microbiology, (Costa et al., 2015) copyright 2015. (B) Graphical depiction of a *V. cholerae* bacterium. Proteins transported by the individual secretion systems are indicated with purple stars, green circles, and red triangles. Substrates of the T2SS and T3SS are translocated through an established secretion system whereas T6SS effectors are located at the tip of the inner tube (1.) and ejected upon extension (2.) and contraction (3.) of the outer sheath. CT = cholera toxin.

1.2.1 Bacterial secretion systems and transporters for transport across one membrane

Proteins spanning the bacterial inner membrane transport proteins across the inner membrane into the periplasm. These proteins might remain in the periplasm, outer

membrane or be delivered into the extracellular space via type II or type V secretion systems in a second step.

The Sec pathway transports unfolded proteins posttranslationally across the inner bacterial membrane (Lycklama and Driessen, 2012). SecB recognizes a newly translated protein by its N-terminal signal sequence of 5–30 hydrophobic amino acids and delivers the protein together with SecA to the SecYEG transmembrane complex (Papanikou et al., 2007). During the subsequent translocation of the protein substrate into the periplasm, the leader sequence is cleaved off (Paetzel et al., 2002). SecA catalyzes the ATP-dependent translocation (Karamanou et al., 1999).

The Tat pathway transports proteins that are folded in the cytoplasm across the inner bacterial membrane (Palmer and Berks, 2012). The Tat signal sequence consists of a twin-arginine consensus motif SRRxFLK (Muller, 2005). The protein complex TatBC recognizes this signal sequence of cytoplasmic proteins and recruits TatA to the complex (Patel et al., 2014). A conformational change of the Tat protein complex triggered by a proton transfer across the inner membrane translocates the substrate protein into the periplasm (Cline et al., 1992). The signal sequence is cleaved off during translocation (Frobel et al., 2012).

Proteins delivered into the periplasm via the Sec pathway can further be delivered into the extracellular space via the T5SS. Proteins transported by this system usually contain multiple domains and are referred to as autotransporters (Leo et al., 2012). One domain of the protein forms a beta-barrel through which it can translocate its remaining domains into the extracellular space. An example of a protein secreted via this mechanism is CdiA of *E. coli* (Aoki et al., 2005) which is involved in bacteria-bacteria interactions and is discussed more in detail in section 1.3.

The T2SS consists of 12–15 proteins and transports folded proteins from the periplasm into the extracellular space (Korotkov et al., 2012). Previous to T2SS-mediated export, proteins are transported into the periplasm via the Sec or Tat pathway (Nivaskumar and Francetic, 2014). Export through the T2SS requires ATP and might require a pseudopilus that remains in the periplasm (Gray et al., 2011). An example of a protein transported by the T2SS of *V. cholerae* is cholera toxin (Sandkvist et al., 2000).

1.2.2 Secretion systems for transport across two membranes

Secretion systems spanning the inner and the outer membranes include the type I, III, IV, and VI secretion systems (Costa et al., 2015). These systems all transport substrates from the cytoplasm directly into the extracellular space or into target cells in one step. (Rego et al., 2010). Substrates transported by these secretion systems are DNA, proteins, or both (Alvarez-Martinez and Christie, 2009; Costa et al., 2015).

The type I secretion system consists of three main components, one in the inner membrane, one in the outer membrane, and one spanning the periplasm (Kanonenberg et al., 2013). The component in the inner membrane is a transporter with an ATP-binding cassette that recognizes a repetitive GGxGxD motif in the C-terminus of protein substrates and mediates their translocation into the extracellular space in an ATP-dependent manner (Hollenstein et al., 2007; Jardetzky, 1966; Jones et al., 2009; Welch, 2001). Examples of proteins secreted by the T1SS are the iron-scavenging protein HasA of *Serratia marcescens* and the pore-forming toxin HlyA of *Escheria coli* which are involved in nutrient acquisition and virulence, respectively (Kanonenberg et al., 2013).

The structure and organization of the T3SS share similarities with the flagellum (Blocker et al., 2003; Desvaux et al., 2006). The whole complex spans both membranes and consists of more than 20 proteins. Chaperones in the bacterial cytoplasm deliver proteins to the secretion system, where a sorting platform organizes their transport directly into eukaryotic cells (Lara-Tejero et al., 2011). T3SS effector proteins like YopE and YopH of *Yersinia pestis* are often associated with host invasion and colonization (Sample et al., 1987).

The T4SS is found in Gram-negative and Gram-positive bacteria (Alvarez-Martinez and Christie, 2009). This system differs from other secretion systems by its ability to transport DNA in addition to proteins (Bradley, 1980; Durrenberger et al., 1991). Translocation by the T4SS is ATP-dependent (Trokter et al., 2014). An example of a T4SS is the Ti system of *Agrobacterium tumefaciens* which mediates the transport of DNA and proteins into plant cells (Alvarez-Martinez and Christie, 2009).

The T6SS contains a molecular puncturing device to translocate proteins into the extracellular space and into the cytoplasm of neighbouring eukaryotic and prokaryotic cells (MacIntyre et al., 2010; Pukatzki et al., 2007; Pukatzki et al., 2006). A contractile

sheath surrounds an inner tube with effector proteins situated at the tip (Basler et al., 2012; Zoued et al., 2014). ATP is required for disassembly of the complex after secretion (Pietrosiuk et al., 2011). Examples of proteins transported by the T6SS are VgrG-1 and VgrG-3 of *V. cholerae* that have antieukaryotic activity and antiprokaryotic activity, respectively (Brooks et al., 2013; Pukatzki et al., 2007). The T6SS is described in more depth in section 1.4.

Taken together, bacteria employ a variety of secretion systems that differ in their secretion machinery and translocate diverse substrates across inner and outer membranes.

1.3 Concept and mechanisms of microbial competition

Bacteria live in a competitive environment and employ a range of mechanisms to compete with neighbouring prokaryotic and eukaryotic cells by directly harming them. Here, I introduce the concept of microbial competition and discuss contact-independent and contact-dependent forms of bacterial competition.

1.3.1 Concept of microbial competition

Bacteria live in diverse environments. Prokaryotic cells are part of an ecosystem that contains a large variety of bacteria, eukaryotic cells, and bacteriophages. Although a genetically diverse group of bacteria is thought to be better prepared for changes, bacteria have evolved selfish mechanisms, such as the T6SS, which work to limit diversity.

Lytic bacteriophages infect bacteria to replicate and induce cell lysis to spread their progeny (Madigan and Martinko, 2006). Temperate phages cause cell lysis in a lytic cycle but also coexist with their bacterial host in a lysogenic cycle (Madigan and Martinko, 2006). The temperate phage genome can be inserted into the bacterial chromosome as prophage. The prophage replicates with the bacterial chromosome. Viral genes are not expressed, unless the phage enters a lytic cycle upon an inducing signal. Bacteriophages impact bacterial populations, for example, *V. cholerae* epidemics have been shown to correlate inversely with the prevalence of lytic vibriophages (Faruque et al., 2005). During infection, intrapatient *V. cholerae* populations likely undergo multiple, independent rounds of selection through phages (Seed et al., 2014). To counteract phages,

V. cholerae modifies the surface antigens required for cell entry of the virus and contains a phage-inducible chromosomal island-like element (PLE) that interferes with phage infection through an unknown mechanism (Seed et al., 2012; Seed et al., 2013). Kim Seed and colleagues described a CRISPR/Cas system (clustered, regularly interspaced short palindromic repeats/CRISPR-associated proteins) in the *V. cholerae*-specific lytic phage ICP1 that allows the virus to interfere with the PLE. CRISPR/Cas systems consist of *cas* genes and a varying number of diverse spacer sequences separated by repeats (Horvath and Barrangou, 2010). Spacers are transcribed into CRISPR RNAs and guide a Cas complex to foreign nucleic acids which are recognized by their complementary sequence to the spacer. The Cas complex subsequently cleaves the targeted nucleic acid strand. When infecting *V. cholerae* with ICP1 phages that lack the CRISPR/Cas spacer specific for the PLE of *V. cholerae*, Kim Seed and colleagues observed an 100,000 fold lower plaquing efficiency (Seed et al., 2013). These observations indicate that, despite interfering with each other, bacteria and phages coexist by using diverse tools to create a balanced equilibrium.

Bacteria face fungal and amoeboid predators in the environment. *Dictyostelium discoideum* is an amoeba that undergoes a multicellular development and preys on bacteria to obtain nutrients (Clarke, 2010). The grazing activity of amoeba was described to control populations of soil bacteria (Clarholm, 1981).

Bacteria compete with each other for space and nutrients and use a variety of mechanisms to win the competition. One form of competition is exploitative competition in which multiple bacteria compete for the same resource and one bacterium indirectly harms another bacterium by taking away its resources. An example of exploitative competition is observed between *Citrobacter rodentium* and *Bacteroides thetaitomicron* during colonization of germ-free mice (Kamada et al., 2012). *C. rodentium* depends on monosaccharides to grow whereas *B. thetaitomicron* can catabolize monosaccharides and polysaccharides. When germ-free mice were colonized with both species and fed a conventional diet of mono- and polysaccharides, over 10^{11} CFUs of *C. rodentium* were recovered per gram of feces (Kamada et al., 2012). After a switch to a simple sugar diet of monosaccharides, 100 fold less CFUs of *C. rodentium* were recovered per gram of feces (Kamada et al., 2012). An additional experiment indicates that this reduction of *C.*

rodentium only occurred in the presence of *B. thetaitomicron*. These data suggest exploitative competition between *C. rodentium* and *B. thetaitomicron* under conditions in which both species depend on the same nutrients for their catabolism (Figure 1-2). Another form of competition is interference competition in which one bacterium actively harms another bacterium in competition (Figure 1-2). Throughout this thesis, I refer to interference competition when using the term competition. In the following sections, I will describe how bacteria use toxins to kill other bacteria in a contact-independent or contact-dependent manner.

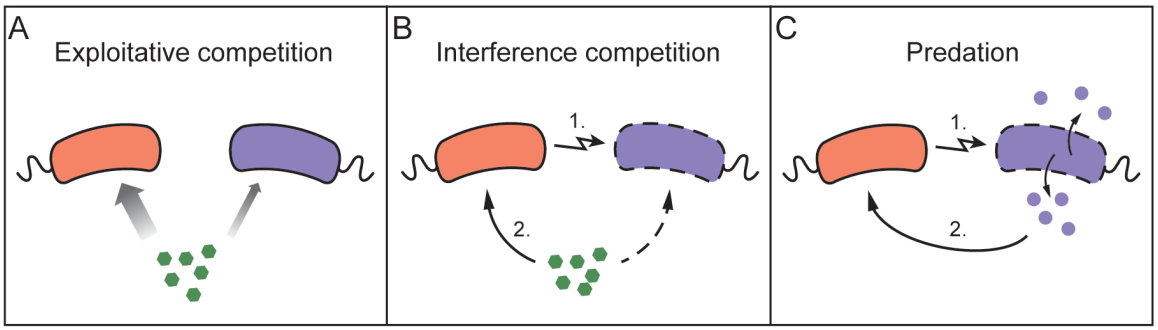


Figure 1-2 Forms of bacterial competition.

(A and B) Examples of competition in which two bacteria compete for the same resource depicted in green. (A) The bacterium depicted in red outcompetes the bacterium depicted in blue by taking away its nutrients. (B) The bacterium depicted in red kills its competitor, indicated by a jagged arrow. The dotted line surrounding indicates a dead bacterium. The intermitted arrow pointing to the dead bacterium indicates that the dead bacterium is not taking up resources any more. Subsequently, the bacterium depicted in red has access to the resource depicted in green. (C) Example of predation between bacteria. The bacterium depicted in red kills the bacterium depicted in blue and subsequently takes up nutrients released by the dead bacterium.

1.3.2 Contact-independent killing

Bacteriocins are proteins that kill bacteria and do not require the producing bacterium to be in direct contact with the target bacterium (Riley, 1998). Gram-negative and Gram-positive bacteria employ bacteriocins (Pugsley, 1984). The toxic activity of bacteriocins can be broadly subdivided into pore-forming toxins, DNAses, RNAses, and peptidoglycan-degrading proteins (Riley, 1998). Immunity proteins specific to their

cognate toxin protect bacteria that produce a bacteriocin from getting killed by its own toxin. Lysis proteins facilitate the release of bacteriocins (Cavard et al., 1985) which require a receptor on the cell surface of the target cell for entry and subsequent deadly action on the target (Lazdunski et al., 1998). Bacteria can become resistant to bacteriocins by modulating the receptor on the cell surface and thus preventing entry of the bacteriocin into the cell (Riley, 1998). The target range of bacteriocins varies from closely related strains to bacteria of other species. The gene encoding a bacteriocin and the gene encoding its cognate immunity protein often come in pairs and can be exchanged by horizontal gene transfer (Rossi et al., 2014).

Bacterial toxins do not need to be proteins. Small compounds like streptomycin form the basis for antibiotics in clinical use and originate from soil bacteria, in this example from *Streptomyces griseus* (Jones et al., 1944).

1.3.3 Contact-dependent killing

Contact-dependent growth inhibition (CDI) and the T6SS are two main mechanisms used by bacteria to kill each other in a contact-dependent manner. S. Aoki and colleagues observed bacteria of the uropathogenic *E. coli* strain EC93 to inhibit the growth of *E. coli* K12 when grown in liquid culture that allowed for direct contact and called this phenomenon "contact-dependent growth inhibition (CDI)" (Aoki et al., 2005). Screening of a cosmid library revealed that the three consecutive genes *cdiB*, *cdiA*, and *cdiI* were sufficient to mediate CDI. The 65 kDa protein CdiB formed a beta barrel in the outer membrane of the bacterium and translocated the 319 kDa CdiA onto the bacterial cell surface. CdiA then formed an extended rod that bound to BamA on the cell surface of a neighbouring bacterium. The C-terminal end of CdiA was then cleaved off and transported into the cytoplasm of the neighbouring cell where it became enzymatically active and inhibited growth by degrading DNA. In the case of uropathogenic *E. coli* strain 536, the toxic C-terminal end of CdiA additionally needs to bind to a permissive factor, the biosynthetic enzyme CysK of the target bacterium, to employ tRNAse activity (Diner et al., 2012). Bacteria can protect themselves from CdiA with the immunity protein CdiI which binds the C-terminus of CdiA and inhibits its toxic activity. Mutations

in BamA that prevent the uptake of CdiA render bacteria resistant to the toxic activity of CdiA.

CDI has been found among alpha, beta, and gamma proteobacteria (Aoki et al., 2010). Whereas most of CdiA is highly conserved, the C-terminal domains and their enzymatic activities differ between strains. C-termini of CdiA with DNAse activity, RNAse activity, and the ability to form pores to interfere with the proton gradient across the bacterial inner membrane of Gram-negative bacteria have been described (Aoki et al., 2010; Aoki et al., 2009). The genome of the *E. coli* strain EC93, for example, contains *cdiA-CT/cdiI* orphan modules that consist only of the 3'end of *cdiA*, which encodes the toxic C-terminus of CdiA, and *cdiI*, which encodes the cognate immunity protein, in addition to a complete set of *cdiB*, *cdiA*, and *cdiI* genes (Poole et al., 2011). Z. Ruhe and colleagues proposed a model in which orphan modules recombine into the complete set of *cdiB*, *cdiA*, and *cdiI* by homologous recombination and allow the bacterium to quickly change the toxic activity of CdiA (Ruhe et al., 2013).

Distribution of very similar *cdi* genes among distantly related bacterial species like *Neisseria lactamica* and *Gallibacterium anatis* suggests the exchange of these genes by horizontal gene transfer (Ruhe et al., 2013). Some genomes encode incomplete 3' end segments of *cdiA* genes in combination with a gene encoding the cognate immunity protein in addition to fully intact *cdiBAI* genes. Z. Ruhe and colleagues proposed a mechanism for the acquisition of these orphan genes to become immune to a larger spectrum of CdiA toxins and a mechanism of reshuffling to equip the intact CdiA stick with a new toxic C-terminal tip (Ruhe et al., 2013).

Microbial competition by the T6SS also depends on direct contact between attacking and target cells. In contrast to toxic C-terminal extensions of CdiA, T6SS effectors are transported in one step into the target cell. As yet, no receptor on the target cell required for T6SS-mediated effector translocation has been described. This difference increases the spectrum of target cells of T6SS-mediated competition. Multiple T6SS effectors can be delivered in one translocation event (Pukatzki et al., 2007). VgrG proteins of the T6SS are similar to CdiA proteins in that they are also structural components of the secretion apparatus that contain various C-terminal extensions with differing enzymatic effector activities (Brooks et al., 2013; Pukatzki et al., 2007). Genes

encoding T6SS effectors are also likely to be exchanged between bacteria by horizontal gene transfer (Unterweger et al., 2014). Bacterial species like *Dicteya dadantii* encode genes for CDI and the T6SS (Aoki et al., 2010; Babujee et al., 2012). Other species like *V. cholerae* encode genes only for the T6SS (Pukatzki et al., 2006).

1.4 The type VI secretion system

Genes encoding proteins of the T6SS have been identified in approximately 25 percent of all Gram-negative bacteria (Boyer et al., 2009). These bacteria belong to the alpha, beta, gamma, and delta proteobacteria and the Bacteriodetes (Ma et al., 2014b; Russell et al., 2014b). Some species like *P. aeruginosa* and *Burkholderia thailandensis* contain multiple T6SSs that target different cells (Mougous et al., 2006; Schwarz et al., 2010). *V. cholerae* contains only one T6SS. Section 1.4.1 provides an overview of T6SS-mediated interactions, and the structure and effectors of this secretion system.

1.4.1 Structure of the T6SS and the mechanism of effector protein ejection

Structural components of the T6SS share structural and functional homology with proteins of the T4 bacteriophage tail (Leiman et al., 2009). The secretion system apparatus is formed of three main components: a base-plate, an inner tube, and a contractile outer sheath. These very basic structural components might vary slightly among the T6SS of individual bacterial species. Bioinformatics analyses revealed three phylogenetically distinct forms of the T6SS (Russell et al., 2014b).

A protein complex anchors the T6SS in the bacterial membrane (Zoued et al., 2014). This complex consists at least of the proteins TssM, TssJ, and TssL (Durand et al., 2012b; Felisberto-Rodrigues et al., 2011; Ma et al., 2009b; VanRheenen et al., 2004). TssM (VasK) is an IcmF family protein localized in the inner membrane and connects the outer membrane protein TssJ with the inner membrane protein TssL. ATP-hydrolysis by TssM facilitates interaction of the protein Hcp with the protein complex of TssM/TssL and is necessary for a functional T6SS (Ma et al., 2009b; Ma et al., 2012). TssF, TssG, and TssA are also ubiquitous in T6SS systems and their role in T6SS function remains to be determined (Boyer et al., 2009; Zheng et al., 2011; Zheng and Leung, 2007). TssK and

TssE are proposed to be part of a baseplate complex that assists in the assembly of the secretion system, similar to the baseplate of contractile bacteriophages (Leiman et al., 2009; Leiman and Shneider, 2012; Zoued et al., 2014). TssK is required for polymerization of the outer T6SS sheath and likely connects the baseplate complex with the membrane anchoring complex through a direct interaction with TssL (Zoued et al., 2013).

The secretion apparatus of the T6SS consists of an inner tube and an outer sheath. Contraction of the device's outer sheath ejects an inner tube from the bacterium into a neighbouring cell (Basler et al., 2012). In the current T6SS model, diverse effector proteins are localized in the conserved tip of the inner tube and are ejected together (Ho et al., 2014; Russell et al., 2014a; Zoued et al., 2014). The inner tube is formed of hexameric rings of the protein Hcp and has an inner diameter of about 40 Å (Mougous et al., 2006). Secretion of Hcp into a culture supernatant is considered the hallmark of a functional T6SS (Pukatzki et al., 2006). A trimer of VgrG proteins is localized in the tip of the Hcp tube (Leiman et al., 2009). VgrG proteins contain at least two domains, one gp27-like and one gp5-like domain (Pukatzki et al., 2007). The very tip of the T6SS is formed by a PAAR domain protein that binds to the tip of the VgrG trimer and forms an inverted cone (Shneider et al., 2013). Shneider and colleagues observed a defect in Hcp secretion in *V. cholerae* mutants deficient in PAAR domain proteins, indicating that these proteins are required for the structural integrity of the T6SS (Shneider et al., 2013).

This inner tube resides within an outer sheath of VipA and VipB. Dimers of these two proteins interact with each other via a four strand beta sheet handshake domain (Kudryashev et al., 2015). Prior to an ejection event, the VipAB sheath, and likely also the Hcp tube, extend across the whole host bacterium. Contraction of the outer sheath results in ejection of the inner tube (Basler et al., 2012). Upon contraction, the VipAB proteins undergo a conformational change and are then recognized by ClpV, an ATPase that recycles the contracted VipAB sheath (Bonemann et al., 2009; Kube et al., 2014).

T6SS effectors are located at the tip of the secretion system and should not universally be considered separately from the secretion apparatus because multiple mutants deficient in T6SS effectors show defects in Hcp secretion (Brooks et al., 2013;

Dong et al., 2013; Miyata et al., 2013; Pukatzki et al., 2007). Thus there is a link between T6SS effectors and the structural integrity of the T6SS.

1.4.2 T6SS effectors

T6SS effectors differ in how they are associated with the tip of the T6SS and in their enzymatic activity which determines the conditions in which killing can be performed and the spectrum of target cells that can be killed (Figure 1-3).

The effectors known to date use five different mechanisms to attach to the tip of the T6SS and are therefore divided into five classes (Ho et al., 2014) (Figure 1-3). Enzymatically active effector domains can be found translationally fused to the C-terminus of VgrG core domains. Examples for VgrG-proteins with C-terminal extensions are VgrG-1 (VC1416) and VgrG-3 (VCA0123) of *V. cholerae* with actin crosslinking activity and peptidoglycan degrading activity, respectively (Brooks et al., 2013; Pukatzki et al., 2007). Many more VgrG proteins have been described with a variety of active domains such as a pe/ppe domain (*B. cenocepacia*), a zinc-metaloprotease domain (*P. aeruginosa*), a mannose binding domain (*Y. intermedia*), a tropomyosin-like domain (*Y. pestis*), a yadA-like and pertactin-like domain (*Y. pestis*), and a yidB-like/fibronectin-like domain (*P. entomophila*) (Pukatzki et al., 2007). The C-terminal extensions are likely cleaved off from the VgrG core protein upon effector delivery into the target cell (Brooks et al., 2013). These effectors are referred to as class 1 effectors.

T6SS effectors can also be encoded in genes in proximity to *vgrG* genes. These effectors are referred to as cargo effectors or class 2 effectors. Examples are TseL (VC1418) and VasX (VCA0020) of *V. cholerae* with lipase activity and pore-forming activity, respectively. The proteins in the VgrG tip with which these cargo effectors interact during translocation need to be determined. Pulldown and immunoprecipitation experiments suggested that cargo effectors directly interact with VgrG proteins (S. Miyata, unpublished observation, and chapter 5). The cargo effectors TseL and VasX are 70 kDa and 120 kDa large, respectively. Some cargo effectors share three 9, 6, and 7 amino acid-long conserved sequences called MIX (marker for type six effectors) (Salomon et al., 2014). The relevance of these conserved sequences has not been

determined. One feature of cargo effectors seems to be their dependence on adaptor proteins ((Miyata et al., 2013); chapter 5). As discussed in detail in chapter 5, we propose that cargo effectors are loaded onto the secretion system with adaptor proteins that themselves are retained in the cytoplasm of the attacking cell.

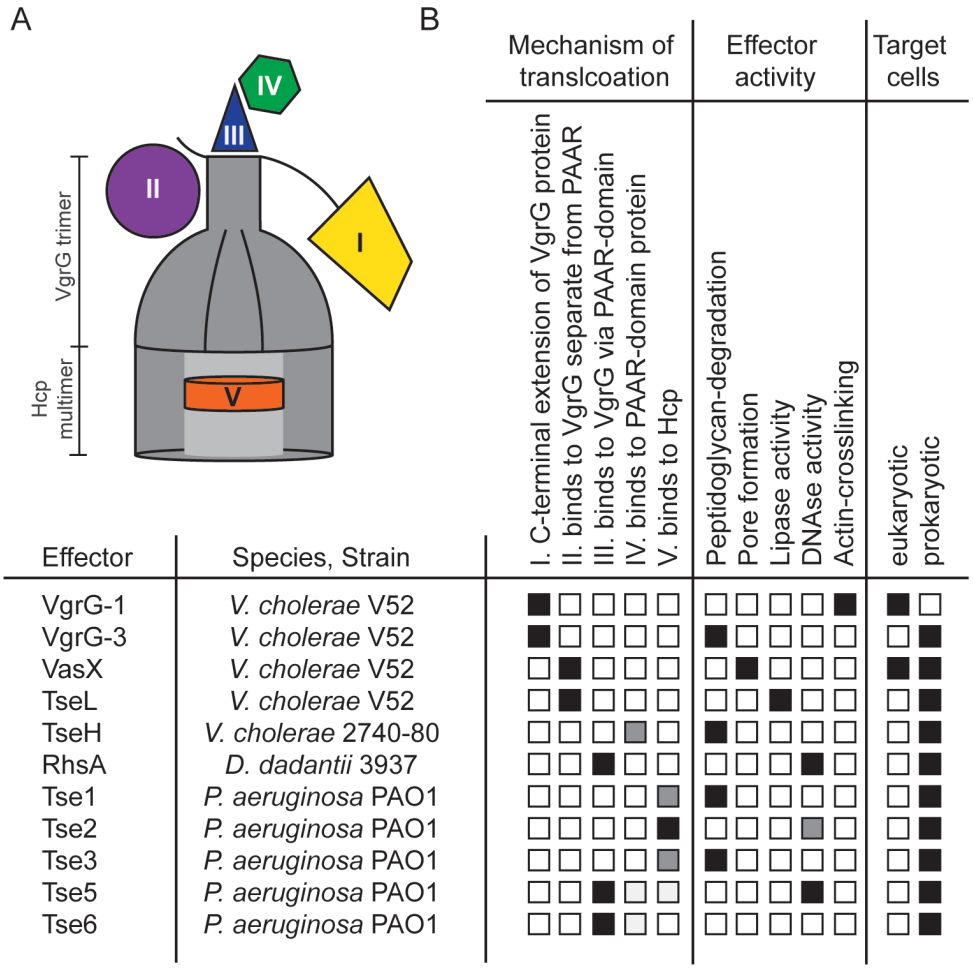


Figure 1-3 Examples of different forms of T6SS effectors.

(A) Graphical depiction of part of the Hcp tube and a VgrG-trimer decorated with a variety of T6SS effectors. One effector for each kind of mechanism with which effectors can be associated with the tip of the secretion system is shown. Explanation of the numbers is provided in the table in B. (B) Overview of the kinds of mechanisms of effector translocation, effector activity, and the target range of effectors. Examples are listed on the left. Black squares indicate that the listed effector has a certain feature. Grey squares indicate that the listed effector putatively has a certain feature but further experimental tests are required. White squares indicate that a listed effector has not (yet) been shown to have an indicated feature.

VgrG proteins contain a motif that mediates a direct interaction with a PAAR domain of a second protein (Shneider et al., 2013). T6SS effectors with PAAR domains are referred to as class 3 effectors; three examples are RhsA of *Dickeya dadantii* with DNAse activity (Koskiniemi et al., 2013; Shneider et al., 2013) and Tse5 and Tse6 of *P. aeruginosa* (Whitney et al., 2014). Shneider and colleagues solved the crystal structure of a PAAR-domain protein with the gp5-like domains of VgrG proteins (Shneider et al., 2013). The resulting structure suggests that PAAR-domain proteins form an inverted cone-like structure at the very tip of the secretion systems. Experiments in *P. aeruginosa* and *Enterobacter cloacae* revealed that PAAR domain proteins depend on cognate VgrG proteins for export (Whitney et al., 2014).

PAAR-domain proteins are also proposed to act as adaptors for T6SS effectors encoded by separate genes (Shneider et al., 2013). These effectors are referred to as class 4 effectors (Ho et al., 2014). An example of such a T6SS effector in *V. cholerae* is TseH with putative peptidoglycan-degrading activity (Altindis et al., 2014). The gene *tseH* is encoded downstream of a gene that encodes a small PAAR-domain protein that has no predicted enzymatic activity. It remains to be tested if TseH depends on the PAAR-domain protein for secretion and if the two proteins directly interact. Some class 3 and 4 effector proteins, including TseH and RhsA, share the characteristic motif GxxxRYxYDxxGRL(I/T) of Rhs proteins (Altindis et al., 2015; Feulner et al., 1990; Koskiniemi et al., 2013).

Small T6SS effectors bind to proteins other than VgrG, that is, to Hcp proteins, and are referred to as class 5 effectors (Ho et al., 2014; Silverman et al., 2013). So far, the best characterized effector in this category is Tse2 of *P. aeruginosa*, with putative nuclease activity (Li et al., 2012; Silverman et al., 2013). The Tse2 effector interacts directly with Hcp, which serves as a chaperone and protects Tse2 from degradation (Silverman et al., 2013). The authors further observed a direct interaction between a class 5 effector and its cognate Hcp protein but not between a class 5 effector and an Hcp protein of a different bacterial species.

Once translocated to the target cell, T6SS effectors employ a variety of enzymatic activities with differing substrate specificities to damage the cell (Figure 1-3). T6SS effectors with different peptidoglycan-degrading activities are widespread among

bacterial species (Russell et al., 2012): amidases cleave carbon-nitrogen bonds, and glycosidases and muramidases hydrolyse glycosidic bonds. Russell and colleagues described amidases with four different substrate specificities and glycosidases with three different substrate specificities among Beta-, Gamma-, and Deltaproteobacteria (Russell et al., 2014a). An example of an amidase is Tse1 of *P. aeruginosa*, an example of a glycosidase is Tse3 of *P. aeruginosa*, and an example of a muramidase is VgrG-3 of *V. cholerae*. The T6SS effectors with peptidoglycan-degrading activity characterized so far belong to classes 1, 3, and 4 T6SS effectors. Bioinformatics predicted class 2 effectors with peptidoglycan-degrading activity (Unterweger et al., 2014) but this prediction remains to be tested experimentally.

Other targets of T6SS effectors are the inner and outer membranes of a bacterial cell. Russell and colleagues described lipases with five different substrate specificities (Russell et al., 2013). An example of a T6SS effector with lipase activity is TseL of *V. cholerae* (Dong et al., 2013). Pore-forming effectors like VasX of *V. cholerae* form a pore and disrupt the electron transport chain along the inner membrane (Miyata et al., 2013). The pore-forming effector VasX is toxic to prokaryotic and eukaryotic cells (Miyata et al., 2011). VasX toxicity depends on an environment of low osmolarity (Miyata et al., 2013) and -mediated killing was not observed in the infant rabbit (Fu et al., 2013). The lipase effectors PldA and PldB have been observed to be toxic to eukaryotic and prokaryotic cells (Jiang et al., 2014). To what extent this is a common feature of these effectors needs to be further examined.

Some effectors target structures in the bacterial cytoplasm. An example of an effector with DNAse activity is RhsA of *Dickeya dadantii (*Koskiniemi *et al., 2013)*. Effectors that were characterized with these activities are among class 3 and class 6 effectors.

Stefan Pukatzki and colleagues characterized the ability of the ACD in the C-terminal extension of VgrG-1 of *V. cholerae* to crosslink actin of murine macrophages and the actin of amoeba (Pukatzki et al., 2007) (described in more detail in 1.5.3).

In summary, T6SS effectors differ in their mechanism of translocation with the T6SS and in their toxic activity (Figure 1-3). Effector targets range from prokaryotic cells (peptidoglycan-degrading effectors) to eukaryotic cells (ACD of VgrG-1) to both cell

types (VasX) (Figure 1-3). The example of the pore-forming effector VasX, for which toxicity depends on an environment of low osmolarity, shows that sometimes the conditions outside the bacterium can affect the toxicity of the effector.

1.4.3 T6SS-mediated cell-cell interactions

Bacteria use the T6SS to mediate interactions with neighbouring prokaryotic or eukaryotic cells. Examples of bacteria that use their T6SS to interact with eukaryotic cells are *V. cholerae*, *Burkholderia thailandensis*, and *P. aeruginosa*.

V. cholerae contains multiple antieukaryotic effectors. The T6SS effector VasX is required for full virulence of *V. cholerae* against the amoeba *Dictyostelium discoideum* (Miyata et al., 2011). The actin-crosslinking domain of *V. cholerae* crosslinks with the actin of amoeba and the actin of murine macrophages (Pukatzki et al., 2007). Amy Ma inoculated infant mice with high titres of *V. cholerae* deficient in the ACD and observed that this mutant colonized infant mice for 16–18h at lower numbers than the wild-type strain (Ma and Mekalanos, 2010).

Burkholderia pseudomallei, *B. mallei*, and *B. thailandensis* use the T6SS to translocate the effector protein VgrG-5 into eukaryotic cells (Schwarz et al., 2014; Toesca et al., 2014). The C-terminal extension of VgrG-5 causes macrophages to form multinucleated giant cells. The enzymatic activity of VgrG-5 remains to be elucidated. *Burkholderia* makes use of these multinucleated, giant cells to spread from cell to cell during infection and requires the T6SS for full virulence in the infection model of mice and hamsters (Pilatz et al., 2006; Schell et al., 2007; Schwarz et al., 2010).

P. aeruginosa uses its T6SS to translocate the two lipases PdlA and PdlB into eukaryotic cells (Jiang et al., 2014). F. Jiang and colleagues observed that a T6SS-deficient mutant of *P. aeruginosa* invaded HeLa cells with 80% lower efficiency than wild-type *P. aeruginosa*. The authors proposed that the lipases act on the PI3K/Akt pathway of eukaryotic cells to facilitate bacterial invasion.

Examples for bacterial species that employ their T6SS for anti-prokaryotic activity are *V. cholerae* and *P. aeruginosa.*

V. cholerae employs at least four antiprokaryotic T6SS effectors (Altindis et al., 2015; Brooks et al., 2013; Dong et al., 2013; Miyata et al., 2011) to kill a variety of Gram-negative bacterial species (MacIntyre et al., 2010). T6SS-mediated killing by *V. cholerae* reduces the number of *E. coli* bacteria it is mixed with from 10^8 to 10^3 within four hours (MacIntyre et al., 2010). *V. cholerae* increases its transformation frequency by killing bacteria when grown on chitin (Borgeaud et al., 2015). Chitinous surfaces render *V. cholerae* naturally competent, a condition in which they take up DNA from the environment (Meibom et al., 2005).

P. aeruginosa uses its T6SS effectors to kill *B. thailandensis, Acinetobacter bayleyi*, *E. coli* and *V. cholerae* in a directional manner (Basler et al., 2013; LeRoux et al., 2015). T6SS-mediated killing by *P. aeruginosa* was described to interfere with the uptake of foreign DNA. *P. aeruginosa* uses its T6SS to kill conjugative *E. coli* and subsequently decreases the number of *P. aeruginosa* conjugants that take up plasmids from *E. coli* by over 5 fold (Ho et al., 2013).

As much as the focus of a toxin is to harm others, bacteria use toxic interactions also to recognize kin and differentiate self from nonself. *Proteus mirabilis* can facilitate the development of kidney stones and is a swarming bacterium that uses its T6SS also for intraspecific competition (Gibbs et al., 2008). Two *P. mirabilis* strains kill each other at the interface of the two swarms in a T6SS-dependent manner and form a boundary between the two swarms that is visible to the naked eye (Alteri et al., 2013).

1.4.4 Regulation and function of the T6SS

Regulation of the T6SS varies from species to species and provides indications for situations in which the T6SS is useful. Regulation of the T6SS in *V. cholerae* and *P. aeruginosa* are examples of the various ways in which the T6SS can be regulated.

P. aeruginosa strains regulate T6SS activity, though an active T6SS under laboratory conditions has been observed for some clinical isolates (Mougous et al., 2006). In general, *P. aeruginosa* activates its T6SS posttranslationally in response to membrane perturbation and danger signals. Among the external stimuli that induce a directional counterattack are conjugative pili and a T6SS-mediated attack of another bacterium. M.

Basler and colleagues mixed *V. cholerae* and *P. aeruginosa* each with a fluorescently tagged VipA to detect T6SS firing events under the fluorescent microscope (Basler et al., 2013). *P. aeruginosa* bacteria did not activate the T6SS until attacked with the T6SS of *V. cholerae* bacteria. *P. aeruginosa* is immune to the T6SS effectors of *V. cholerae* and survives the attack. The *V. cholerae* attack is sensed by the *P. aeruginosa* TagQRST/PpkA signaling cascade which phosphorylates Fha1 and induces a directional T6SS-mediated counterattack, and *V. cholerae* succumbs to the counterattack it originally initiated. In a population of *P. aeruginosa*, *V. cholera* with an active T6SS, and *V. cholerae* with an inactive T6SS, *P. aeruginosa* kill *V. cholerae* with an active T6SS whereas *V. cholerae* with an inactive T6SS remain untouched (Basler et al., 2013). The ability of *P. aeruginosa* to induce a directional counterattack might be used to eliminate certain bacteria from a population.

LeRoux and colleagues recently studied how *P. aeruginosa* responds when mixed with *B. thailandensis* that kill *P. aeruginosa* bacteria in a T6SS-dependent manner (LeRoux et al., 2015). The authors found that *P. aeruginosa* that dies as a result of a T6SS-mediated attack of *B. thailandensis* releases a diffusible signal. This compound is sensed by other *P. aeruginosa* bacteria which in turn posttranscriptionally activate their T6SS to outcompete *B. thailandensis*. Taken together, these observations suggest that an individual *P. aeruginosa* bacterium activates its T6SS in response to being attacked and in response to dying kin bacteria.

1.4.5 The T6SS and bacterial pathogenesis

The T6SS is found in bacterial species that are pathogenic to humans and plants. Examples are *V. cholerae*, *P. aeruginosa*, *Burkholderia pseudomallei, Yersinia pseudotuberculosis*, and *Dickey dadantii*.

V. cholerae is a human pathogen that induces severe diarrhoea and causes up to 120,000 deaths around the globe annually (WHO, 2014c). The T6SS is also found in nonpandemic *V. cholerae* strains even though a correlation between pathogenic T6SS effectors and pandemic strains has been observed (Unterweger et al., 2014). Detection of T6SS gene expression in vivo (Mandlik et al., 2011), the activating effect of the host

factor mucin on the T6SS (V. Bachmann et al., unpublished observation), and the role of the ACD in *V. cholerae* colonization of the infant mouse (Ma and Mekalanos, 2010) suggest a role of the T6SS in vivo. The role of the T6SS during infection is discussed in detail in chapter 7 of this thesis. Several bacteria that use a T6SS for survival and competition are described below.

P. aeruginosa infects the lungs and is the major cause of morbidity and mortality of cystic fibrosis (CF) patients (Govan and Deretic, 1996). Antibodies against the Hcp of the T6SS from *P. aeruginosa* were detected in the blood and Hcp was detected in the mucus of CF patients (Mougous et al., 2006). These observations suggest T6SS activity during infection although the role of the T6SS during infection and to what extent the T6SS is required for a successful infection remain to be elucidated.

B. pseudomallei causes pneumonia and abscesses with a mortality rate of up to 40% (Wiersinga et al., 2012). The T6SS has been shown to be required for giant cell formation. The ability of bacteria to replicate intracellularly and to invade the host through direct cell-to-cell spread is considered crucial for the pathogenicity of *B. pseudomallei*.

Y. pseudotuberculosis causes abdominal pain, diarrhoea, vomiting, and fever (PAHO, 2001). Outbreaks of illness caused by *Y. pseudotuberculosis* have been noted; for example, in Japan, 732 people became ill and 134 had to be hospitalized in 1991 (PAHO, 2001). The T6SS of *Y. pseudotuberculosis* secretes an effector with the ability to bind Zn^{2+} ions (Wang et al., 2015). Mice infected with a *Y. pseudotuberculosis* mutant deficient in the T6SS effector and also deficient in a Zn^{2+} transporter all survived (Wang et al., 2015), whereas only one quarter of mice infected with wild-type *Y. pseudotuberculosis* were still alive 20 h after infection (Wang et al., 2015).

D. dadantii is a plant pathogen that infects a wide range of plants including tomatoes, sweet potatoes, carrots, and onions (Mansfield et al., 2012). At least two T6SS effectors have been proposed, but the role of the T6SS in *D. dadantii* pathogenesis remains to be characterized.

Recently, the T6SS has been discovered in nonpathogenic bacterial species. An example is *Bacteroides fragilis*, a member of the intestinal microbiota (Russell et al., 2014b). Expression of T6SS genes was observed by quantitative RT-PCR in germ-free

mice colonized with *B. fragilis* for one week (Russell et al., 2014b). The ability of *B. fragilis* to kill another member of the microbiota, *B. thetaitomicron*, was determined in vitro in a killing assay (Russell et al., 2014b).

1.5 The type VI secretion system of *V. cholerae*

V. cholerae is among the bacterial species with the currently best studied T6SS. The following sections describe what is currently known about the T6SS gene clusters and T6SS effectors of *V. cholerae*, regulation of the T6SS, and interactions between *V. cholerae* and other cells mediated by this secretion system.

1.5.1 The T6SS enables *V. cholerae* to kill eukaryotic and prokaryotic cells

V. cholerae uses its T6SS to interact with eukaryotic and prokaryotic cells (Figure 1-4). Stefan Pukatzki performed a very elegant screen of transposon mutants to determine the virulence factors of the *V. cholerae* strain V52 that kill the amoeba *Dictyostelium discoideum* and subsequently prevent their predation on bacteria (Pukatzki et al., 2006). This screen resulted in the identification of transposon mutants that lost their virulence against *D. discoideum* as a result of transposon insertions in genes of the T6SS gene clusters. These findings set the ground for a model in which *V. cholerae* uses the T6SS to interact with neighbouring cells such as amoeba. The characterization of the T6SS effector VgrG-1, which is required for virulence against *D. discoideum*, was further characterized to mediate cytotoxicity to murine macrophages and thus increased the target spectrum of the T6SS from amoeba to endocytic, eukaryotic host cells (Ma et al., 2009a; Pukatzki et al., 2007).

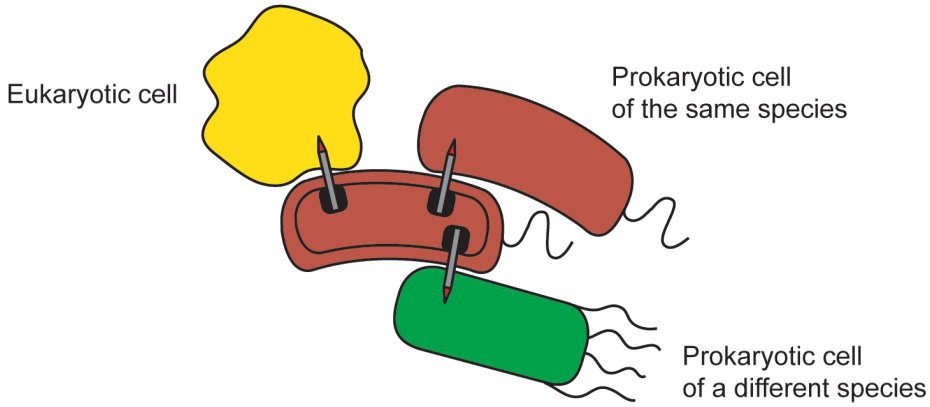


Figure 1-4 Overview of the T6SS-mediated interactions of *V. cholerae.*

A *V. cholerae* bacterium is depicted in the middle; T6SSs are coloured in black and grey to indicate interactions with eukaryotic cells (depicted in yellow, top left), other prokaryotic cells of the species *V. cholerae* (depicted in red, top right) and prokaryotic cells of a different species (depicted in green, bottom right).

Dana MacIntyre and colleagues showed that *V. cholerae* uses its T6SS to interact with other prokaryotic cells (MacIntyre et al., 2010) (Figure 1-4). For example, after a 4 hour incubation of 10^8 *V. cholerae* bacteria with 10^7 *E. coli* bacteria, only about 10^2 surviving *E. coli* were recovered. About 10^8 surviving *E. coli* were recovered after incubation with a *vasK*-deficient *V. cholerae* mutant that does not have an active T6SS, indicating that killing depends on an active T6SS. Other Gram-negative bacterial species—*Salmonella enterica* serovar Typhimurium, *Citrobacter rodentium*, enterohemorrhagic *E. coli* and enteropathogenic *E. coli*—were also subject to T6SS-mediated killing by *V. cholerae* (MacIntyre et al., 2010). *V. cholerae* strains N16961 and C6706 did not die in the presence of the *V. cholerae* strain V52 with an active T6SS and seemed to be immune to T6SS-mediated killing. The Gram-positive bacteria *Enterococcus faecilis*, *Bacillus subtilis*, *Listeria monocytogenes*, and *Staphylococcus aureus* also did not die in response to a T6SS-mediated attack. Separation of *V. cholerae* and *E. coli* by a filter that prevents direct contact between the two species but not the exchange of diffusible molecules abrogated T6SS-mediated killing suggesting its dependency on direct contact. Although T6SS-mediated cytotoxicity toward eukaryotic cells depends on the ACD of VgrG-1, killing of *E. coli* does not. The T6SS effectors VgrG-3, VasX, TseL, and TseH have antiprokaryotic activity (Altindis et al., 2015;

Brooks et al., 2013; Dong et al., 2013; Miyata et al., 2013). T6SS-mediated killing of *V. cholerae* has been observed not only on LB agar plates but also in the intestine of the infant rabbit (Fu et al., 2013) and on chitin (Borgeaud et al., 2015).

Together, these experiments established a model in which *V. cholerae* uses its T6SS to interact with amoeba, host macrophages, and other bacteria. Eukaryotic and prokaryotic cells die in response to a T6SS-mediated attack. The molecular mechanism underlying the T6SS-mediated killing of bacteria is discussed in chapter 2 of this thesis.

1.5.2 The T6SS gene clusters

Proteins of the *V. cholerae* T6SS are encoded by 36 genes organized in three gene clusters and one separate effector module distributed over the small and large chromosome of *V. cholerae* (Altindis et al., 2015; Pukatzki et al., 2006) (Figure 1-5). We do not yet call the clusters operons because an exhaustive search for promoters by 5' extension, and mRNA transcript mapping by northern blot and the PCR have not been performed. The genome of the *V. cholerae* strain N16961 was sequenced in 2000 (Heidelberg et al., 2000) and is used as a reference genome to investigate the T6SS gene clusters. The large T6SS gene cluster contains genes encoding structural and regulatory proteins of the T6SS and is located on the large chromosome. The proteins VipA and VipB form the contractile outer sheath (Basler et al., 2012), VasH, which is a regulator protein (Kitaoka et al., 2011a), and VasK (TssM), which is an essential inner membrane protein (Pukatzki et al., 2006), are encoded by genes in the large T6SS gene cluster. The genes *vgrG-3* and *tsiV3* are located at the 3' end of the T6SS gene cluster and encode a peptidoglycan-degrading effector and its cognate immunity protein, respectively.

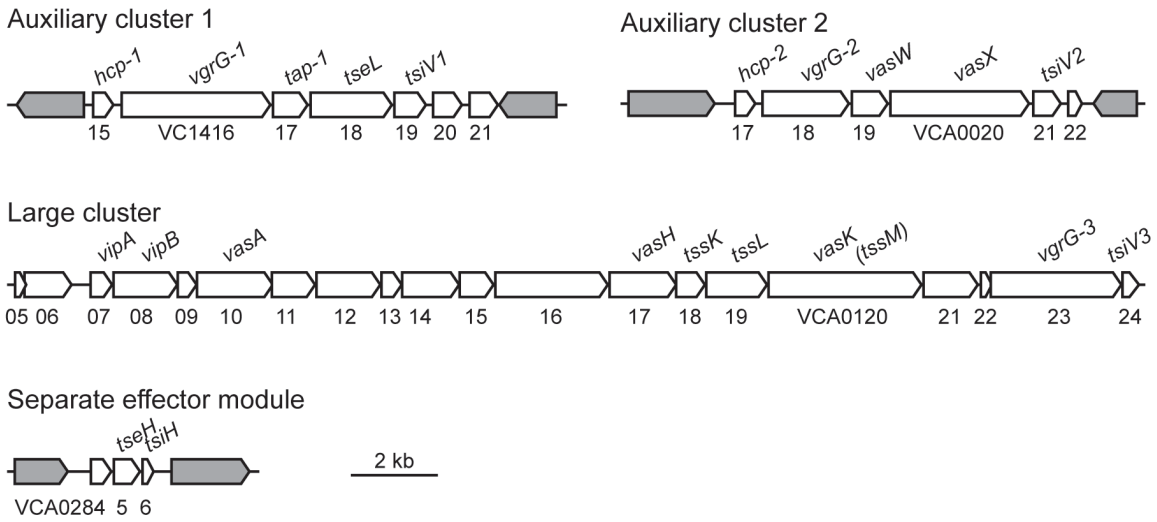


Figure 1-5 T6SS gene clusters of *V. cholerae* strain N16961.

Graphical depiction of the T6SS gene clusters (drawn to scale). Genes of the T6SS clusters are indicated in white, surrounding genes are coloured in grey. No surrounding genes of the large cluster are shown for space reasons. The names are indicated on top and the locus tags are indicated below the respective genes. Genes with the locus tag VCA0105 and VCA0284 encode PAAR-domain proteins.

In the *V. cholerae* strain N16961 gene clusters, auxiliary clusters 1 and 2 are located on the large and small chromosome, respectively, and share similarities in the kind and order of genes they contain (Figure 1-5). The first gene in cluster 1 is *hcp-1* and the first gene in cluster 2 is *hcp-2*; these genes encode identical proteins and the proteins they encode form the inner sheath of the T6SS. The genes downstream of *hcp-1* and *hcp-2* encode VgrG-1 and VgrG-2 proteins, respectively. The three following genes *tap-1*, *tseL*, and *tsiV1* in auxiliary cluster 1 and *vasW*, *vasX*, and *tsiV2* in auxiliary cluster 2 encode adaptor proteins, cargo effectors, and cognate immunity proteins, respectively (Dong et al., 2013; Miyata et al., 2013). A varying number of open reading frames are encoded downstream of *tsiV1* and *tsiV2* that encode putative orphan immunity proteins (Appendix A, Figure 9-6).

The T6SS effector TseH is encoded in a gene on the small chromosome outside of the T6SS gene cluster (Figure 1-5) (Altindis et al., 2015). The gene upstream of *tseH* encodes a PAAR-domain protein and the gene downstream of *tseH* encodes the immunity protein TsiH. A second PAAR-domain protein is encoded at the very 5'end of the large T6SS cluster in the gene with the locus tag VCA0105 (Shneider et al., 2013).

1.5.3 T6SS effectors of *V. cholerae*

Five T6SS effectors with antiprokaryotic and antieukaryotic activities have been described in *V. cholerae* strains N16961, V52, and 2740-80 (Table 1).

VgrG-1 contains a C-terminal extension with an actin-crosslinking domain (ACD) (Table 1) (Pukatzki et al., 2007). Pukatzki and colleagues added recombinant VgrG-1 to cell lysates of *D. discoideum* and murine macrophages and observed actin-crosslinking in both cell types. *V. cholerae* without an ACD loses its virulence toward *D. discoideum* (Ma et al., 2009a). *V. cholerae* with the ACD kills the amoeba and no amoeba predation was observed. Amy Ma and colleagues also added *V. cholerae* to cells of the murine macrophage cell line J774 and observed the release of lactate dehydrogenase (LDH) as a marker for cytotoxicity. *V. cholerae* with an ACD caused over 20% more LDH release compared to ACD-deficient *V. cholerae* indicating that the ACD in the *V. cholerae* VgrG-1 promotes cytotoxic activity. In another experiment, Amy Ma and colleagues tested if VgrG-1 is translocated into eukaryotic cells in a T6SS-dependent manner. The authors translationally fused *blaM*, which encodes the beta lactamase enzyme Bla, to VgrG-1 and loaded murine macrophages with CCF2/AM, which is a FRET-based substrate for Bla and emits light of a different wavelength when catalyzed by Bla (Ma et al., 2009a). Macrophages incubated with *V. cholerae* with an active T6SS changed their emission wavelength indicating that VgrG-1 reached the macrophage cytoplasm and catalyzed CCF2/AM whereas *V. cholerae* mutants deficient in a functional T6SS did not. Actin-crosslinking occurred only among phagocytic cells. When *V. cholerae* was added to Chinese ovary hamster cells that express Fcγ receptors to induce opsonophagocytosis upon IgG stimulation, actin-crosslinking was observed. This antieukaryotic activity of the ACD is required for *V. cholerae* to maintain colonization of the infant mouse at high titres (Ma and Mekalanos, 2010). Amy Ma infected infant mice intragastrically with *V. cholerae* and determinedthe CFU recovered from the small intestine 2 hours and 16–18 hours after inoculation. *V. cholerae* with an ACD was recovered at numbers equal to those of a mutant without the ACD at the 2-hour timepoint but at about 1,000-fold higher numbers at the later time-point. Therefore, VgrG-1 is translocated into eukaryotic cells and contains an ACD that crosslinks actin and is cytotoxic to amoeba and macrophages. The *V. cholerae* genome contains two genes, *vgrG-1* and *rtxA*, that encode ACDs. In the

experiments described above, a *V. cholerae* strain deficient in *rtxA* was used to analyze the activity of the ACD of VgrG-1 (Sheahan et al., 2004). Sheahan and colleagues speculated that one of the proteins acquired an ACD by gene duplication. The advantage of having an ACD encoded in two separate proteins has not been determined. Durand and colleagues determined the crystal structure of the VgrG-1 ACD (Durand et al., 2012).

VgrG-3 contains a C-terminal extension with peptidoglycan-degrading activity (Table 1-1) (Brooks et al., 2013). Teresa Brooks and colleagues subjected recombinant VgrG-3 and the C-terminal extension only to gel electrophoresis on a 12% SDS acrylamide gel supplemented with peptidoglycan. In-gel peptidoglycan was stained with methylene blue and zones of clearing at the expected sizes of full length VgrG-3 and the C-terminal extension indicated their abilities to degrade peptidoglycan (Brooks et al., 2013).

VasX is a cargo effector with antieukaryotic and antiprokaryotic activities (Miyata et al., 2011; Miyata et al., 2013). Sarah Miyata and colleagues observed that VasX was secreted into culture supernatants in a T6SS-dependent manner and was required for full virulence against the amoeba *D. dictyostelium* (Miyata et al., 2011). Growth of the amoeba was observed on a lawn of VasX-deficient *V. cholerae* but not on a lawn of wild-type *V. cholerae*. Sarah Miyata and colleagues also showed that VasX binds lipids and that VasX interferes with the integrity of the inner cell membrane, which explains its toxic activity against prokaryotic and eukaryotic cells (Miyata et al., 2011; Miyata et al., 2013).

Effector	Mechanism of translocation	Effector activity	Target cells	First described by
VgrG-1	Part of VgrG	Actin-crosslinking	Eukaryotic	Pukatzki *et al.*, 2007
VgrG-3	Part of VgrG	Peptidoglycan-degradation	Prokaryotic	Brooks *et al.*, 2013
VasX	Cargo effector	Pore-formation	Pro- & Eukaryotic	Miyata *et al.*, 2011
TseL	Cargo effector	Lipase activity	Prokaryotic	Dong *et al.*, 2013
TseH	Not determined	Peptidoglycan-degradation	Prokaryotic	Altinidis *et al.*, 2015

Table 1-1 Overview of T6SS effectors shared by *V. cholerae* strains V52, 2740-80, and N16961.

TseL is another cargo effector with antiprokaryotic activity (Dong et al., 2013; Russell et al., 2013). Tao Dong and colleagues mapped the catalytically active site of

TseL to an aspartic acid at position 425 (Dong et al., 2013). Alistair Russell and colleagues determined the substrate specificity of TseL by adding recombinant protein to vesicles of phospholipid derivatives with fluorescent moieties at various stereochemical positions. The results indicated that TseL cleaves the ester bond in phospholipids.

TseH is the most recently described T6SS effector with antiprokaryotic activity (Altindis et al., 2015). When TseH with a periplasmic leader sequence was expressed in *E. coli*, E. Altinids and colleagues recovered 10,000 fold less *E. coli* bacteria compared to *E. coli* in which the expression of periplasmic TseH was repressed with glucose. The results of a bioinformatics analysis suggested that TseH has peptidoglycan degrading activity. This hypothesis was further supported by an experiment in which recombinant TseH at similar levels to lysozyme caused cell lysis of *E. coli* with a permeabilized outer membrane. E. Altindis and colleagues detected TseH only in culture supernatants of bacteria with an active T6SS, suggesting that TseH is translocated with the T6SS. The mechanism by which TseH is associated with the T6SS secretion system remains to be elucidated.

The protein encoded in the gene with the locus tag VCA0118 was proposed by Zheng and colleagues to be a T6SS effector (Zheng et al., 2011). The authors drew this conclusion based on reduced levels of *E. coli* killing despite Hcp secretion of a *V. cholerae* mutant deficient in VCA0118. This putative effector has not yet been further studied.

In summary, *V. cholerae* encodes T6SS effectors with antieukaryotic and antiprokaryotic activity. The mechanisms by which these effectors are associated with the T6SS categorize them as class 1 and 2 effectors. TseH might be a class 4 effector. No effectors with class 3 and class 5 characteristics have yet been described for *V. cholerae*. As described in chapter 2 of this thesis, *V. cholerae* strains harbor immunity proteins against prokaryotic effectors to protect from a sister-cell mediated attack. Chapter 4 describes *V. cholerae* strains that harbor a variety of effectors beyond the ones shared by the strains V52, 2740-80, and N16961 described in this section.

1.5.4 Regulation of the *V. cholerae* T6SS

Individual *V. cholerae* strains regulate their T6SSs differently. The strain V52 has an active T6SS under laboratory conditions whereas the pandemic strain C6706 represses its T6SS under these conditions.

Deciphering the regulation of genes within the T6SS gene cluster is an ongoing effort. VasH is encoded in a gene in the large T6SS gene cluster and is an activator of the alternative sigma factor 54 (Kitaoka et al., 2011a). Maya Kitaoka and colleagues observed that VasH induces gene expression of *hcp2*, *vgrG-2*, and *vasX* in the auxiliary cluster 2. Putative sigma factor 54 binding sites were predicted upstream of auxiliary clusters 1 and 2 (Williams et al., 1996). Electrophoretic mobility shift assays revealed that VasH binds to promoters upstream of *hcp-1* and *hcp-2* in the auxiliary clusters (Bernard et al., 2011) and supports a model of VasH-regulated expression of genes in the auxiliary clusters (Figure 1-6 A).

The T6SS gene clusters have been shown to be activated in strains with a repressed T6SS under laboratory conditions by the regulatory pathways of quorum sensing and natural competence (Figure 1-6 A, B). Conditions that mimic high cell density have been shown to induce T6SS gene expression (Borgeaud et al., 2015; Shao and Bassler, 2014; Zheng et al., 2011). Small regulatory RNAs contribute to the regulation of the T6SS by repressing HapR and by base pairing (Shao and Bassler, 2014). Strains deficient in *tsrA* in addition to a lack of *luxO* also activate their T6SS posttranslationally, as indicated by Hcp secretion into culture supernatants. Strains with a functional HapR additionally require the expression of Tfox to activate the transcriptional regulator QstR and induce expression of T6SS genes (Borgeaud et al., 2015).

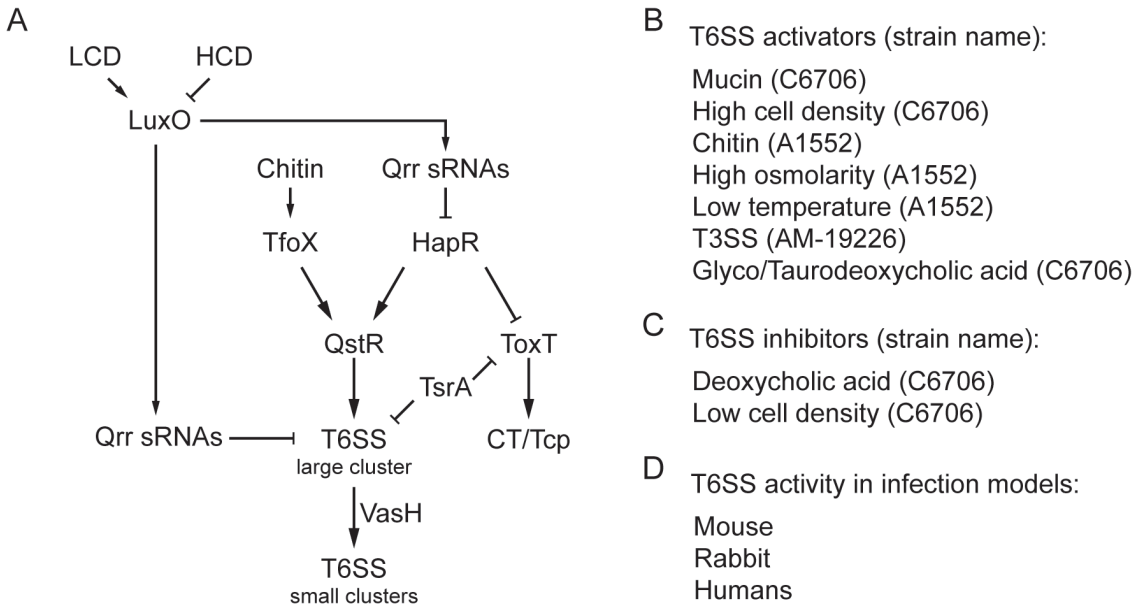


Figure 1-6 Overview of T6SS regulation in *V. cholerae.*

(A) Regulation of the T6SS by regulators and small RNAs. The model was combined from (Borgeaud et al., 2015; Shao and Bassler, 2014; Zheng et al., 2010). HCD: high cell density; LCD: low cell density. (B-D) Overview of activators and inhibitors of T6SS activity. The strain name that was used in the respective experiments is indicated. List of infection models in which T6SS activity was observed. See the text for details and references.

External stimuli (other than chitin and high cell density) that result in an active *V. cholerae* T6SS are high osmolarity, low temperature (Ishikawa et al., 2012), and host mucin (V. Bachmann et al., under review) (Figure 1-6 A, B). Studies on the gene expression profile of *V. cholerae* indicate that the T6SS of pandemic strains is activated during infection of humans and animals (Lombardo et al., 2007; Mandlik et al., 2011) (Figure 1-6 D).

Regulation of the T6SS and the T3SS is linked in strains that have both secretion systems (Chaand and Dziejman, 2013) (Figure 1-6 B). Proteins of the T3SS are encoded in genes on a 50 kb-long genomic island. Two genes on this island encode the regulators $VttR_A$ and $VttR_B$ that induce expression of T3SS genes in a bile-dependent manner (Alam et al., 2010). Comparison of genome-wide gene expression via RNA-Seq of samples from wild-type O37 serogroup strain AM-19226 and a *vttr*$_A$-deficient mutant

grown on bile revealed 3–5 fold lower expression of the T6SS genes *hcp-1*, *hcp-2*, and *vgrG-1* in the absence of *vttrA* (Chaand and Dziejman, 2013).

1.5.5 *V. cholerae* strains and their use to characterize the T6SS

Multiple *V. cholerae* strains, such as N16961, C6706, V52 and A1552, were used to characterize the T6SS activity of *V. cholerae*. N16961 is a *V. cholerae* strain derived from a clinical isolate obtained in 1975 in Bangladesh. N16961 belongs to the O1 serogroup and El Tor biotype. It is associated with the 7th *V. cholerae* pandemic and was the first strain to be fully sequenced and to have an annotated genome (Heidelberg et al., 2000). The strain is not frequently used anymore because of the recent identification of a frameshift mutation in *hapR*, the gene encoding a regulator protein that controls a variety of cellular processes such as virulence, motility, and biofilm formation (Zhu et al., 2002). The genome of the strain N16961 and its annotation are still used as a reference.

V. cholerae strain C6706 is another strain associated with the 7th pandemic and derived from a clinical isolate obtained in 1991 in Peru. This strain does not have an active T6SS under laboratory conditions. It is immune to a T6SS-mediated attack of strain V52 and has been used to study T6SS regulation and T6SS-mediated bacterial killing in vivo (Fu et al., 2013; Zheng et al., 2011). Unlike the C6706 isolate used in the original description of the natural competence of *V. cholerae* (Meibom et al., 2005), our C6706 isolate is not naturally competent when grown on chitin.

V. cholerae strain V52 belongs to the O37 serogroup. It is derived from a rectal swab of a cholera patient who died in a refugee camp with a contaminated water well in Sudan during a local outbreak in 1968 (Zinnaka and Carpenter, 1972). This strain has an active T6SS under laboratory conditions and is therefore used for most of our experiments to study the T6SS and its effects on eukaryotic and prokaryotic cells (MacIntyre et al., 2010; Pukatzki et al., 2007; Pukatzki et al., 2006). Like strain C6706, strain V52 is not naturally competent when grown on chitin. Comparison of the V52 genome sequence to that of other *V. cholerae* strains indicates that V52 is closely related to O1 serogroup strains (Chun et al., 2009). Among the O1 serogroup strains, V52 shares more sequence similarities with classical biotype than with El Tor biotype strains (Chun

et al., 2009). However, V52 is distinct from O1 serogroup strains of the classical biotype (Chun et al., 2009).

V. cholerae strain A1552 belongs to the O1 serogroup and the El Tor biotype. This strain becomes naturally competent when grown on chitin. The T6SS is not active when grown under laboratory conditions until, in the presence of chitin, genes that encode components of the T6SS and of the DNA uptake machinery are activated. Using this strain in their experiments, Sandrine Borgeaud and colleagues observed an increased transformation frequency upon T6SS-mediated killing on chitin surfaces (Borgeaud et al., 2015).

The T6SS gene clusters are very similar among the four strains N16961, C6706, V52 and A1552. It is expected that the T6SS gene cluster of the strain N16961 encodes one additional open reading frame (VCA0122) that encodes a protein (that has not been characterized) with putative regulatory function (Zheng et al., 2011) and that is not present in the T6SS gene cluster of the strain V52.

1.6 Hypothesis and aims

Cholera is a global health burden that has caused millions of deaths in the past centuries and still maintains a pandemic spread in the 21st century (WHO, 2014c). The etiological agent of cholera is the bacterium *V. cholerae*, which is abundant in the marine environment.

The *V. cholerae* species and the environments in which these bacteria are found are diverse (Chun et al., 2009). In the environment, *V. cholerae* bacteria live associated with phylo- and zooplankton in biofilm communities (Tamplin et al., 1990; Teschler et al., 2015). Within these communities, it is expected that bacteria interact with each other and with bacteria of other species. Competitive behavior among various bacteria, and between bacteria and amoebae, arises because of limited nutrients and space in the environment (Hibbing et al., 2010). *V. cholerae* strains of the O1 serogroup are capable of causing pandemic cholera. During infection of the human host, *V. cholerae* bacteria interact with each other, and with the microbiota and phagocytic eukaryotic cells of the

immune system. The microbiota and the immune system provide barriers to pathogen infections.

The type VI secretion system enables *V. cholerae* to kill eukaryotic and prokaryotic cells (MacIntyre et al., 2010; Pukatzki et al., 2007; Pukatzki et al., 2006). The secretion system forms a contractile sheath around an inner tube that translocates effector proteins associated with its tip into neighbouring cells. Effectors are toxic unless they are bound by an immunity protein and thus inactivated (Basler et al., 2012; Brooks et al., 2013; Dong et al., 2013; Ma et al., 2009a). Based on studies of the role of the *V. cholerae* type VI secretion system in pathogenesis, the Pukatzki laboratory studies the role of the *V. cholerae* type VI secretion system in pathogenesis and develops new therapeutic strategies by targeting the T6SS.

I started this PhD project just before Dana MacIntyre and colleagues described how *V. cholerae* uses its T6SS to kill prokaryotic cells (MacIntyre et al., 2010). The experiments revealed how the *V. cholerae* strain V52 kills *E. coli* in a T6SS-dependent manner whereas the *V. cholerae* strain C6706 was immune to a T6SS-mediated attack (MacIntyre et al., 2010). At that time, the molecular mechanism of T6SS-mediated killing was unknown and the big unanswered question was how *V. cholerae* kills other bacteria. I hypothesized that *V. cholerae* strains contain T6SS effectors that kill bacteria like *E. coli* on contact and immunity proteins that protect bacteria like the *V. cholerae* strain C6706 from a T6SS-mediated attack. I screened transposon mutants of the strain C6706 and identified three T6SS immunity proteins TsiV1, TsiV2, and TsiV3 that protect from the cognate effectors TseL, VasX, and VgrG-3. I furthermore asked if T6SS-dependent competition occurs among environmental *V. cholerae* isolates from the Rio Grande Delta. Subsequent experiments revealed T6SS-mediated killing between *V. cholerae* strains. We then wondered about the molecular basis for intraspecific T6SS-mediated competition and performed bioinformatics analyses of multiple *V. cholerae* strains. These analyses revealed effector modules that contain effector and immunity protein-encoding genes. These modules differ from otherwise conserved genes in the gene clusters and are likely exchanged between strains by horizontal gene transfer. It was still unknown how diverse T6SS effectors without an obvious signal sequence are translocated with a conserved secretion system. I then identified a diverse group of chimeric adaptor proteins that are

required for the secretion of their cognate effector. Finally, pandemic strains contain the same set of T6SS effector modules. Because the strains that caused the 6th cholera pandemic were not yet analyzed regarding their competitive fitness, I asked how the strains that caused the 6th pandemic used their T6SS to resist a T6SS-mediated attack. My experiments suggested that *V. cholerae* strains that caused the 6th pandemic were weak microbial competitors.

The results of these experiments contribute to our understanding of how the T6SS of *V. cholerae* kills other bacteria and provide new insights into the diversity of T6SS effector and immunity proteins within the species *V. cholerae* (Figure 1-7). The observation that pathogenic *V. cholerae* strains share the same T6SS effector and immunity proteins, and as a result can coexist among each other but not with nonpandemic strains, forms a basis to translate these findings into applications with a direct benefit for human health.

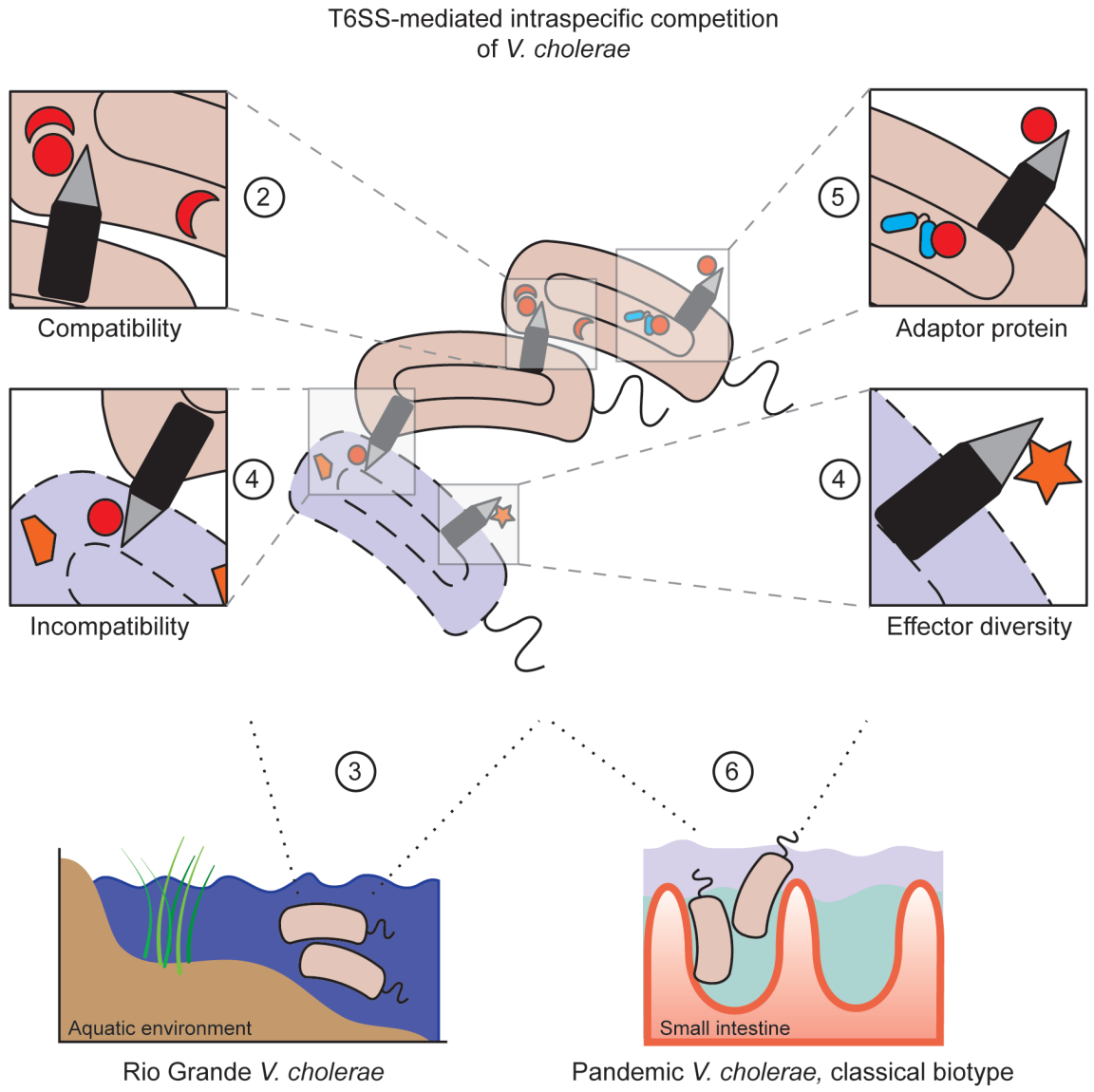


Figure 1-7 Overview over the major findings described in this thesis.

Graphical depiction of three *V. cholerae* bacteria. Two bacteria depicted in red are able to coexist. One of the bacteria depicted in red kills the bacterium depicted in blue in a T6SS-dependent manner. The numbers refer to chapters in this thesis where the depicted activity is explained.

Chapter 2

Identification of T6SS immunity protein-encoding genes *tsiV1*, *tsiV2*, and *tsiV3*

Portions of this chapter have been published as:

Miyata, S.T., Unterweger, D., Rudko, S. and Pukatzki, S., Dual expression profile of type VI secretion system immunity genes protects pandemic *Vibrio cholerae*. *PLoS Pathogens*, 2013. 9(12): p. e1003752.

Unterweger, D., Miyata, S.T., Bachmann, V., Mullins, T., Brooks, T., Kostiuk, B., Provenzano D. and Pukatzki, S., The *Vibrio cholerae* type VI secretion system employs diverse effector modules for intraspecific competition. *Nature Communications*, 2014. 5:3549.

Figure 1-4 B: The killing assay was performed by Sarah Miyata and is published in Miyata et al., 2013 (see above).

2 Identification of T6SS immunity protein-encoding genes *tsiV1*, *tsiV2*, and *tsiV3*

2.1 Introduction

Vibrio cholerae employs its T6SS to harm prokaryotic and eukaryotic cells (MacIntyre et al., 2010; Pukatzki et al., 2007; Pukatzki et al., 2006). The antibacterial properties of the *V. cholerae* T6SS were first observed against *E. coli* (MacIntyre et al., 2010). An extended analysis showed that *V. cholerae* employs its T6SS also to kill other Gram-negative bacteria such as *Salmonella* Typhimurium, *Citrobacter rodentium*, enterohemorrhagic *E. coli*, and enterophathogenic *E. coli* (MacIntyre et al., 2010). The *V. cholerae* O37 strain V52, isolated from a cholera patient in a refugee camp in Sudan in 1968 (Zinnaka and Carpenter, 1972), was used in these competition experiments because it employs an active T6SS under laboratory conditions. The pandemic O1 serogroup strain C6706 represses its T6SS under laboratory conditions and is immune to a T6SS-mediated attack by V52 (MacIntyre et al., 2010; Zheng et al., 2010).

The *V. cholerae* T6SS delivers effectors to neighboring cells and corresponding antagonistic proteins confer immunity against effectors that are delivered by sister cells. The T6SS effectors that *V. cholerae* uses to kill prokaryotic cells were unknown when this phenomenon was first described in 2010 (MacIntyre et al., 2010). The only characterized *V. cholerae* T6SS effector at this time was VgrG-1, which was not necessary for *E. coli* killing. VgrG-1, a T6SS effector protein with an actin-crosslinking domain (ACD), mediates killing of murine macrophages (Pukatzki et al., 2007). Although VgrG-1 is not needed for killing *E. coli* (MacIntyre et al., 2010), we envisioned that antiprokaryotic effectors might be transported with the T6SS into a target bacterium, act toxic, and thus kill the bacterium. We speculated that *V. cholerae* strains immune to a T6SS-mediated attack might contain immunity proteins that protect from the toxic activity of antiprokaryotic T6SS effectors.

This chapter describes how I made use of the observation that *V. cholera* of the strain C6706 are immune to a T6SS-mediated attack of *V. cholera* of the strain V52. I used a transposon library of C6706 (Cameron et al., 2008) to identify mutants that had lost their protection to a T6SS-mediated attack due to a transposon-insertion in an immunity protein-encoding gene. In a second step, I identified genes encoding candidates

for antiprokaryotic T6SS effector proteins based on their location in proximity to immunity protein-encoding genes.

2.2 Results

2.2.1 Transposon insertions in three T6SS genes results in lost immunity of C6706 to a T6SS-mediated attack of V52

V. cholerae strain C6706 has previously been shown to be immune to a T6SS-mediated attack of the *V. cholerae* strain V52 (MacIntyre et al., 2010). We hypothesized that *V. cholerae* strains such as C6706 encode immunity proteins within the T6SS gene clusters that are expressed independently of the other genes in the T6SS gene cluster, protecting the bacterium from a T6SS-mediated attack by an active T6SS of a neighbouring bacterium.

To identify such immunity genes, I used a transposon library of C6706 with transposon insertions in nonessential genes, including 30 mutants with single insertions in T6SS genes (Cameron et al., 2008). In a killing assay, we mixed 10^7 bacteria of the prey strain C6706 with 10^8 bacteria of the predator strain *V. cholerae* V52. Bacteria of the strain C6706 were also mixed with V52Δ*vasK*, a mutant that does not engage in T6SS-mediated killing because VasK is required for a functional T6SS. Bacterial mixtures were incubated on LB agar plates for four hours at 37 ºC. Spots were harvested, serially diluted, and plated onto selective plates to determine the number of surviving V52 and C6706. A competitive index was determined by calculating the ratio of the number of colony forming units of a surviving C6706 mutant exposed to wild-type V52 over the number of colony forming units of a surviving C6706 mutant exposed to V52Δ*vasK*. *E. coli* was exposed to V52 and V52Δ*vasK* to control for comparable experimental conditions. The negative index of *E. coli* of ~ 10^{-3} to 10^{-4} shows that this experimental setup allows for similar levels of T6SS-mediated killing than previously observed (MacIntyre et al., 2010). When comparing the competitive indices of all the transposon mutants, mutants with insertions in the genes VCA0021 (*tsiV2*), VCA0124 (*tsiV3*), and VC1418 (*tseL*) (Figure 2-7) showed the lowest competitive indices, which differed by over 100-fold with the competitive index of parental C6706. Mutants with transposon

insertions in genes with the locus tag VCA0114, VCA0115, and VCA0121 were also susceptible to killing by V52. However, we did not pursue these three genes further because these genes can be deleted in V52 (Zheng et al., 2011). It is therefore unlikely that genes with the locus tag VCA0114, VCA0115, and VCA0121 encode immunity proteins because V52's constitutive T6SS would have selected against null-mutations.

Previous attempts to delete *tsiV2* (S. Miyata, unpublished observation) and *tsiV3* (Zheng et al., 2011) in V52 were unsuccessful, indicating that these two genes might be required for survival. Thus, we decided to focus on *tsiV2*, *tsiV3*, and *tseL* and determine the role of these genes and their products in conferring immunity of the strain C6706 to a T6SS-mediated attack of the strain V52.

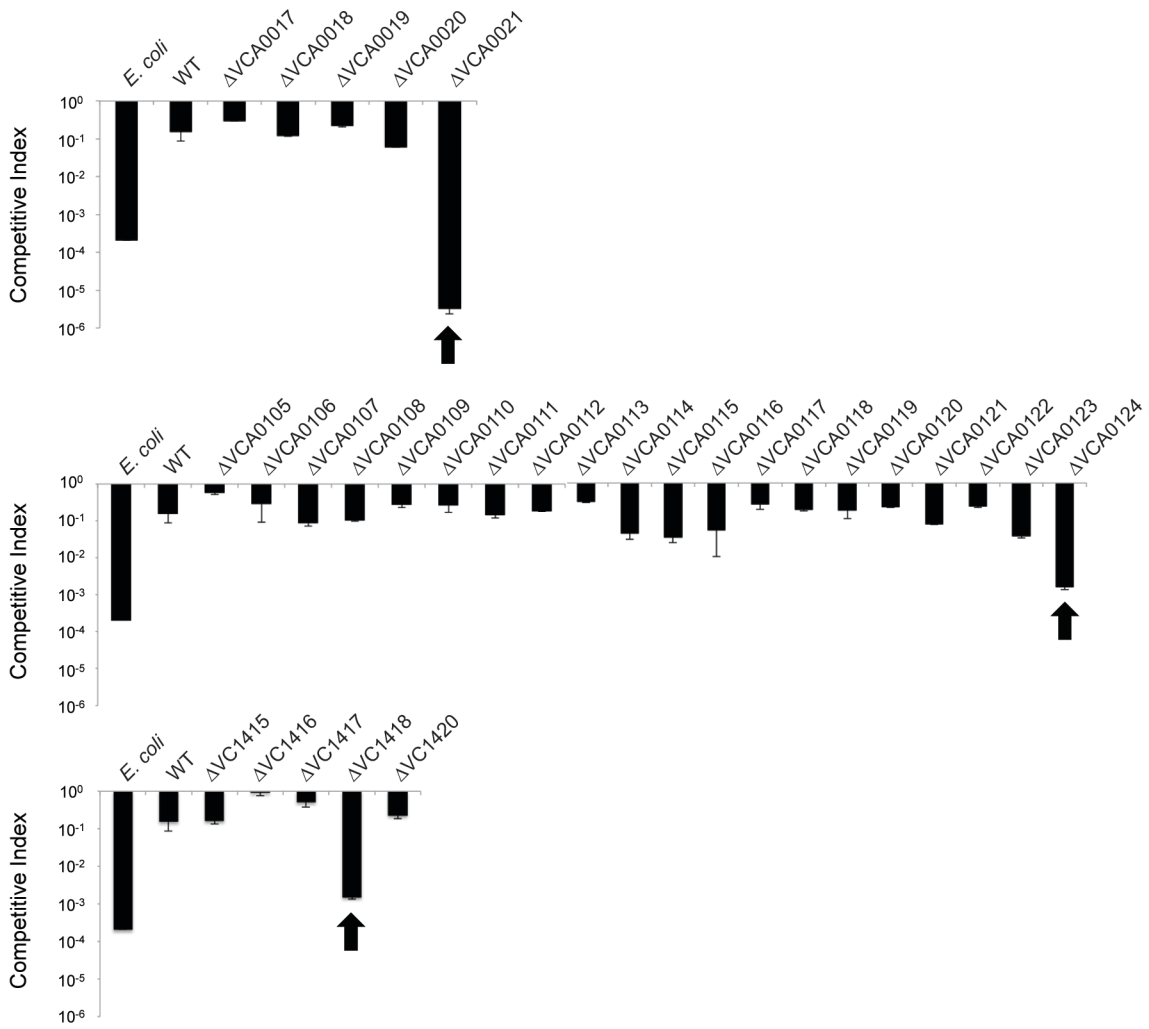


Figure 2-1 Identification of T6SS immunity protein-encoding genes in *V. cholerae*.

Killing assays were performed to screen a C6706 T6SS transposon library for mutants that became sensitive to killing by *V. cholerae* V52. Predator strains included wild-type V52 and the T6SS-null strain V52Δ*vasK* (negative control). *E. coli* strain MG1655 was included as a positive control as this strain had previously been shown to be susceptible to killing by V52. The data are presented in three individual graphs, each representing one of the *V. cholerae* T6SS gene clusters (i.e., VCA0017-VCA0021 – top panel, VCA0105-VCA0124 – middle panel, and VC1415-VC1420 – bottom panel). The competitive index was calculated by dividing recovered CFU after exposure to V52 by recovered CFU exposed to V52Δ*vasK*. Arrows indicate C6706 mutants identified as sensitive to killing by V52. These data represent three independent experiments performed in technical duplicate. Error bars indicate the standard deviation.

2.2.2 Episomal expression of TsiV1 – not TseL – restores survival of C6706Δ*tseL*

To validate the role of *tseL* (VC1418) in mediating immunity of C6706 to a T6SS-mediated attack of V52, we hypothesized that *tseL* expression in trans would restore immunity of an in-frame *tseL* deletion mutant. In-frame gene deletion mutants were created by allelic exchange (Metcalf et al., 1996). To express the immunity-protein encoding genes *in trans* upon induction with arabinose, we PCR-amplified *tseL* and ligated a restricted DNA fragment into plasmid pBAD24 downstream of an arabinose-inducible promoter.

A killing assay was performed in which the *tseL* in-frame deletion mutant of C6706 was exposed to wild-type V52 or V52Δ*vasK*. The number of C6706 surviving a 4 hour-exposure to the indicated V52 strain was plotted (Figure 2-2). In the absence of arabinose, C6706Δ*tseL* harbouring the *tseL*-containing plasmid was killed by wild-type V52 but not V52Δ*vasK*. This observation confirmed the results we previously obtained in our screen of transposon mutants. To test if we could rescue this mutant by inducing the expression of *tseL*, we exposed complemented C6706Δ*tseL* to V52 and V52Δ*vasK* under inducing conditions. Unexpectedly, induction of *tseL* expression did not protect C6706Δ*tseL* exposed to wild-type V52. This finding suggested that lost immunity of C6706Δ*tseL* is observed not because of the lack of the protein TseL but rather because of a polar effect on downstream genes.

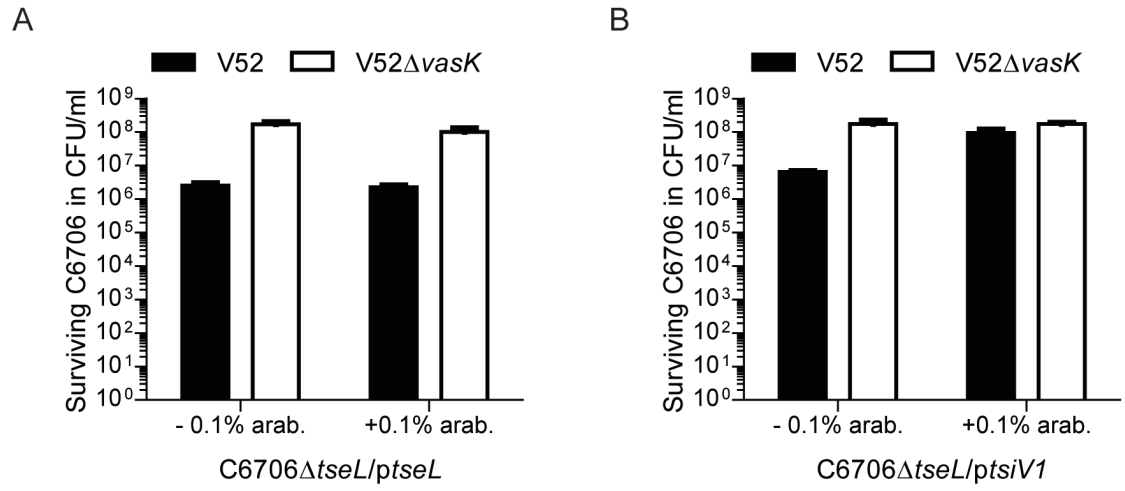


Figure 2-2 Episomal expression of *tsiV1*, not *tseL* restores immunity of C6706Δ*tseL*.

Survival of rifampicin-resistant C6706Δ*tseL* harboring either (A) pBAD24-*tseL*::FLAG, or (B) pBAD24-*tsiV1*::FLAG was determined by measuring CFU following exposure to the indicated rifampicin-sensitive predator V52 or V52Δ*vasK* in a killing assay. These data were obtained from one representative experiment performed in technical duplicate. Error bars indicate the standard deviation.

To test if episomal expression of *tsiV1*, the gene encoded downstream of *tseL*, restores immunity, we exposed C6706Δ*tseL*/p*tsiV1* to V52 in a killing assay. Induction of *tsiV1* expression increased the survival rate of C6706Δ*tseL* by > 10-fold compared to complementation with an empty vector or episomal *tseL*. This indicated to us that the nucleotide sequence of *tseL* likely contains an internal promoter that drives *tsiV1* expression and provides protection from TseL.

2.2.3 Complementation of *tsiV1*, *tsiV2*, and *tsiV3* in-frame deletion mutants restores immunity to a T6SS-mediated attack

To further validate the role of TsiV1, TsiV2, and TsiV3 in protecting C6706 from a T6SS-mediated attack by V52, we hypothesized that the lost immunity of a *tsiV* in-frame deletion mutant could be restored by expressing the respective *tsiV in trans*. Therefore, we created in-frame deletions of *tsiV1*, *tsiV2*, or *tsiV3* in C6706. All three immunity genes were then cloned into pBAD24 to complement in-frame deletions.

A killing assay was performed in which the C6706 mutants were exposed to wild-type V52 or V52Δ*vasK*. C6706 mutants lacking *tsiV1*, *tsiV2*, or *tsiV3* transformed with empty vector were recovered at > 100-fold lower numbers than parental C6706. This observation confirmed the results of the original screen in which transposon mutants of *tsiV2* and *tsiV3* became sensitive to a T6SS-mediated attack of V52 (Figure 2-1) and further suggested that *tsiV1* encodes an immunity protein.

Expression of the immunity genes in trans restored the recovery of about 10^7 survivors of C6706 similar to the surviving numbers of parental C6706. This was observed for *tsiV1*, *tsiV2*, and *tsiV3*. From this observation, we concluded that the three proteins TsiV1, TsiV2, and TsiV3 mediate immunity to a T6SS-mediated attack of V52.

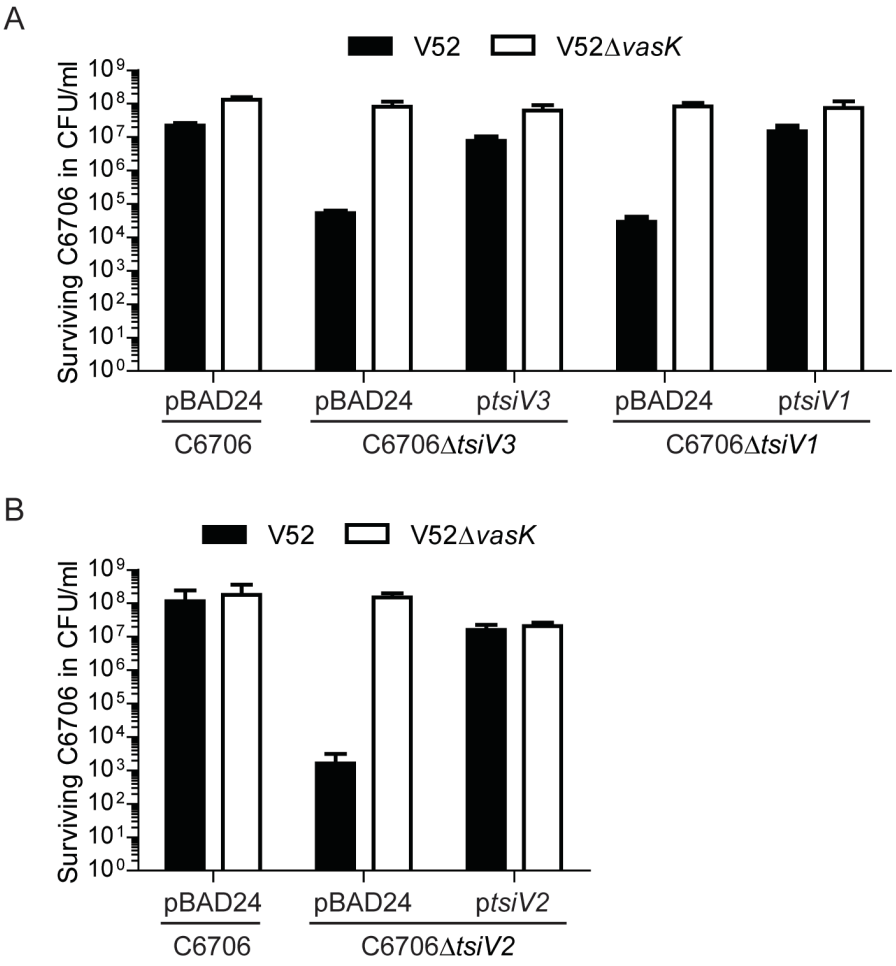


Figure 2-3 TsiV1, TsiV2, and TsiV3 mediate immunity to a T6SS-mediated attack of V52.

Episomal *tsiV1*, *tsiV3*, and *tsiV2* protect C6706Δ*tsiV1*, C6706Δ*tsiV3*, and C6706Δ*tsiV2*, respectively, from V52. Survival of rifampicin-resistant prey strains C6706, C6706Δ*tsiV3*, C6706Δ*tsiV1*, or C6706Δ*tsiV2* harboring an empty vector (pBAD24), pBAD24-*tsiV3*::FLAG, pBAD24-*tsiV1*::FLAG, or pBAD24-*tsiV2*::FLAG was determined by measuring CFU following exposure to a rifampicin-sensitive predator (listed in the legend) – wild-type V52 or V52Δ*vasK* that serves as a negative control for T6SS-mediated bacterial killing. Arabinose was included in all samples to drive expression from the PBAD promoter. These data represent two independent experiments, each performed in technical duplicate. Error bars indicate the standard deviation.

2.2.4 Immunity protein-encoding genes require the gene upstream in cis to mediate protection

We previously observed that the loss of immunity of a *tseL*-deletion mutant of C6706 can be restored by providing *tsiV1* and not *tseL in trans* (Section 2.2.2) indicating that the gene upstream of *tsiV1* is important for establishing immunity in C6706. To test if the requirement of the gene upstream of the *tsiV*s is a general phenomenon that also applies to *tsiV2* and *tsiV3*, I created in-frame deletion mutants in C6706. The gene encoded upstream of *tsiV2*, *vasX*, and the gene encoded upstream of *tsiV3*, *vgrG-3*, were deleted. The resulting mutants were complemented in trans with the deleted gene or the *tsiV* gene encoded downstream and tested in a killing assay.

C6706 deficient in *vgrG-3* provided with an empty vector was recovered at about 10-fold lower numbers after exposure to V52 compared to V52Δ*vasK* (Figure 2-4 A). This result indicated that *vgrG-3* is required for the protection of C6706 from a T6SS-mediated attack of V52. Further complementation showed that *vgrG-3* needs to be provided in cis of *tsiV3* to protect C6706. A similar result was obtained when analyzing the *vasX*-deficient mutant of C6706 (Figure 2-4 B).

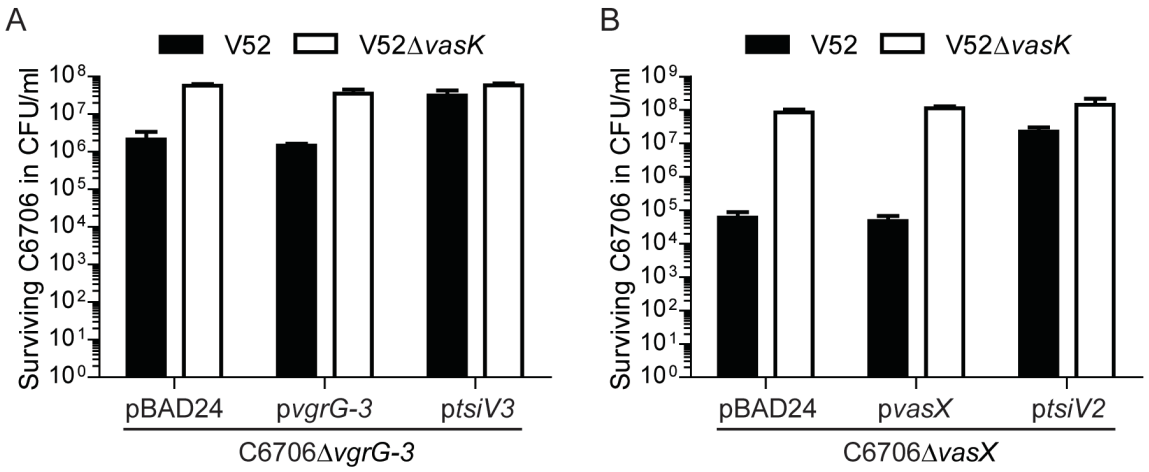


Figure 2-4 C6706 deletion mutants lacking the full length of *vgrG-3* or *vasX* lose immunity toward T6SS-mediated killing.

Survival of rifampicin-resistant (A) C6706Δ*vgrG-3* or (B) C6706Δ*vasX* and complemented strains (indicated on the x-axis) was determined by measuring CFU following exposure to the indicated rifampicin-sensitive predator listed in the legend. These data represent three independent experiments performed in technical duplicate. Error bars indicate the standard deviation.

These results indicate that the nucleotide sequences of the genes upstream of the *tsiV*s are required for mediating immunity to a T6SS-mediated attack. The mechanism behind the T6SS-mediated attack still needed to be identified at that time.

2.2.5 Immunity proteins protect against cognate effectors

Proteins with similar protective functions, such as TsiV1, TsiV2, and TsiV3, occur in several T6SS-independent bacterial toxin-systems, for example, the contact-dependent growth inhibition (CDI) mechanism of *E. coli* (Aoki et al., 2005). CdiI protects from the toxic function of CdiA, encoded in a gene directly upstream of *cdiI*. The T6SS gene clusters, including the *tsiV*s and their surrounding genes, are also found in the strain V52 which is capable of killing C6706 derivatives lacking individual immunity protein-encoding genes. Closer analysis of the genes surrounding the *tsiV*s in V52 concentrated on the genes *tseL*, *vasX*, and *vgrG-3* that are encoded immediately upstream of *tsiV1*, *tsiV2*, and *tsiV3*, respectively. TseL and TsiV1 exemplify two main indicators of a toxin-antitoxin pair with TseL as the putative effector protein and TsiV1 as the cognate

immunity protein. First, the open reading frames of *tseL* and *tsiV1* overlap by 4 basepairs. Overlaps between the open-reading frames of toxin- and antitoxin-encoding genes are described for various toxin-antitoxin systems. For example, *ldrX* and *ldrE* of a type I toxin-antitoxin system in *E. coli*; *relB* and *relE* of a type II toxin-antitoxin system in *Acinetobacter baumanii* have short overlaps (Brantl, 2012; Jurenaite et al., 2013). Second, the identification (by BLAST) of a motif characteristic for lipases suggested that TseL may function as a putative antiprokaryotic effector. Similar indicators for toxic effector proteins were found for VasX and VgrG-3. This led us to hypothesize that TsiV1, TsiV2, and TsiV3 protect C6706 from TseL, VasX, and VgrG-3, respectively. To test this hypothesis, I created in-frame deletion mutants V52Δ*vasX*Δ*vgrG-3*, V52Δ*tseL*Δ*vgrG-3*, and V52Δ*tseL*Δ*vasX* that each encode only one of the three genes *tseL*, *vasX*, or *vgrG-3*. Subsequently, I exposed C6706 mutants that encode one or none of the three genes *tsiV1*, *tsiV2*, or *tsiV3* to the V52 mutants and tested which immunity protein-encoding gene protects C6706 from V52 equipped with one of the three effectors.

First, wild-type C6706, C6706Δ*tsiV1*Δ*tsiV2*Δ*tsiV3*, and the three C6706 mutants with one immunity protein-encoding gene each were exposed to wild-type V52 (Figure 2-5 left panel). The lack of *tsiV1-3* reduced the numbers of surviving C6706 by > 10,000 fold. C6706Δ*tsiV2*Δ*tsiV3* or C6706Δ*tsiV1*Δ*tsiV2* were killed at levels similar to that of C6706Δ*tsiV1*Δ*tsiV2*Δ*tsiV3*. C6706Δ*tsiV1*Δ*tsiV3* was recovered at about 10 fold higher numbers than C6706Δ*tsiV1*Δ*tsiV2*Δ*tsiV3*. These results indicated that full protection of C6706 from a V52-mediated attack requires the presence of multiple *tsiV* genes and that C6706 mutants encoding only *tsiV1* or *tsiV3* are less protected than the mutant that only encodes *tsiV2*.

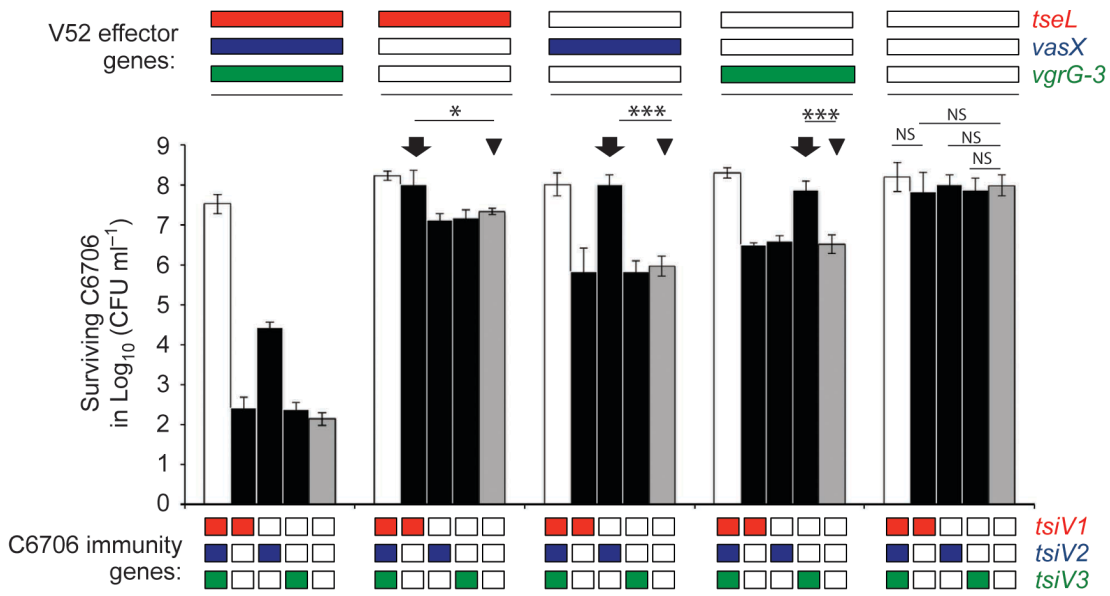


Figure 2-5 TsiV1, TsiV2, and TsiV3 protect from TseL, VasX, and VgrG-3, respectively. A competition assay of V52 effector mutants (top) against immunity mutants of C6706 (bottom) at a 10:1 ratio was performed. The logarithm of the number of surviving wild-type or mutant C6706 colony-forming units is shown on the y axis. White, black, and grey bars represent C6706 containing all three, one, or no immunity protein-encoding genes, respectively. In-frame gene deletions are indicated by an empty rectangle. Triangles indicate toxic activity of individual T6SS effectors. Arrows highlight protection from the effector by its cognate immunity gene. Log-transformed data of two independent experiments, each performed in duplicate, are the basis for the arithmetic mean and the s.d. shown as bars and error bars, respectively. Stars indicate statistical significance (unpaired, two-tailed Student's t-test: *P<0.05; ***P<0.0005; NS, P>0.05). Reprinted by permission from Macmillan Publishers Ltd: Nature Communications, Unterweger et al., 2014, copyright 2014.

Next, we exposed the same set of C6706 mutants and wild-type C6706 to V52 lacking *vasX* and *vgrG-3* but encoding *tseL*. A C6706 mutant lacking all three *tsiV*s survived by 10 fold less than wild type C6706, indicating that one of the *tsiV*s is mediating immunity to this V52 mutant. Comparing the number of surviving C6706 mutants encoding one of three *tsiV*s, only the presence of *tsiV1* restored immunity to V52 comparable to parental C6706 whereas the presence of *tsiV2* or *tsiV3* did not show any protective effect. This result suggests that TsiV1 protects from a TseL-mediated attack.

In a similar way, the indicated C6706 mutants were exposed to V52 encoding *vasX* but not *tseL* or *vgrG-3*. C6706 lacking all *tsiV*s was recovered about 2-fold less than C6706 wild-type, indicating that a lack of *tsiV*s renders C6706 susceptible to V52-mediated killing. When comparing the C6706 mutants encoding one immunity gene at a time, the mutant encoding *tsiV2* was recovered at a number similar to that of C6706 wild-type, whereas mutants encoding *tsiV1* or *tsiV3* were not. This observation suggests that TsiV2 inhibits VasX-mediated killing.

Last, wild-type C6706 or its immunity mutants were exposed to V52 encoding only *vgrG-3*. C6706 deficient in all three *tsiV*s was recovered by over 10 fold less than C6706 wild-type. C6706 encoding *tsiV3* but not *tsiV1* or *tsiV2* was recovered at a level similar to that of C6706 wild-type. This finding indicated that TsiV3 specifically inhibits VgrG-3.

C6706 mutants exposed to V52 lacking *tseL*, *vasX*, and *vgrG-3* were all recovered at about 10^8 colony forming units per killing spot, similar to C6706 wild-type. This suggests that the previously observed killing of C6706 mutants lacking all three *tsiV*s was mediated based on the presence of *tseL*, *vasX*, or *vgrG-3*.

In summary, we tested if the TsiVs protect from killing mediated by TseL, VasX or VgrG-3. My results suggest that immunity proteins protect against their cognate effectors encoded upstream.

2.2.6 TseL, VasX, and VgrG-3 are required for killing of *E. coli* and maintaining structural integrity of the T6SS

To further explore the role of *tseL*, *vasX*, and *vgrG-3* for T6SS-mediated killing, I tested their requirement for V52 to kill *E. coli*. At the time, the best characterized T6SS effector was VgrG-1 with antieukaryotic activity. In addition to its role as an effector, VgrG-1 was also known to contribute to the structural integrity of the T6SS, because a *vgrG-1* deletion mutant of V52 secreted less Hcp than wild-type V52 (Pukatzki et al., 2007). To test if TseL, VasX, and VgrG-3 also contribute to the structural integrity of the T6SS, I analyzed the Hcp secretion profile of deletion mutants lacking *tseL*, *vasX*, and *vgrG-3*.

In-frame single gene deletion mutants of the genes *tseL*, *vasX*, and *vgrG-3* were created in strain V52 in addition to the mutants V52Δ*tseL*Δ*vasX*, V52Δ*tseL*Δ*vgrG-3*, V52Δ*vasX*Δ*vgrG-3* and V52Δ*tseL*Δ*vasX*Δ*vgrG-3* I created for previous experiments (see section 2.2.5). All of these mutants were tested for their ability to kill *E. coli* in a killing assay.

Wild-type V52 kills *E. coli* whereas V52Δ*vasK* does not, as shown previously (MacIntyre et al., 2010) (Figure 2-6). Mutants of V52 lacking either *vasX* or *tseL* killed *E. coli* at levels similar to that of wild-type V52, whereas a *vgrG-3*-deficient mutant killed *E. coli* at a 10-fold lower level than that of wild-type V52 (Figure 2-6). The double deletion mutants lacking the two genes *vasX* and *tseL* killed at higher levels than the double mutants in which *vgrG-3* is knocked out in combination with either *vasX* or *tseL*. When exposed to the triple deletion mutant, V52Δ*tseL*Δ*vasX*Δ*vgrG-3*, *E. coli* survived at a number similar to *E. coli* survival after exposure to V52Δ*vasK*, indicating that no killing occurred upon deletion of *tseL*, *vasX*, and *vgrG-3*.

To investigate whether the genes *tseL*, *vasX*, and *vgrG-3* contribute to a functional T6SS, I analyzed Hcp secretion. Hcp forms the inner tube of the secretion system and its detection in culture supernatant is considered the hallmark of a functional T6SS (Pukatzki et al., 2006).

I collected the pellets and supernatants of wild-type V52, V52Δ*vasK*, V52Δ*vasX*, V52Δ*vgrG-3*, V52Δ*tseL*, V52Δ*vasX*Δ*vgrG-3*, V52Δ*vasX*Δ*tseL*, V52Δ*tseL*Δ*vgrG-3* and V52Δ*vasX*Δ*vgrG-3*Δ*tseL*. The proteins of the individual samples were separated by size using SDS polyacrylamide gel electrophoresis (SDS PAGE) and blotted onto nitrocellulose membranes for detection with antibodies. DnaK is a cytoplasmic protein that was detected as a loading control in the pellet and as a control for the detection of proteins in the supernatant due to cell lysis and inactive secretion.

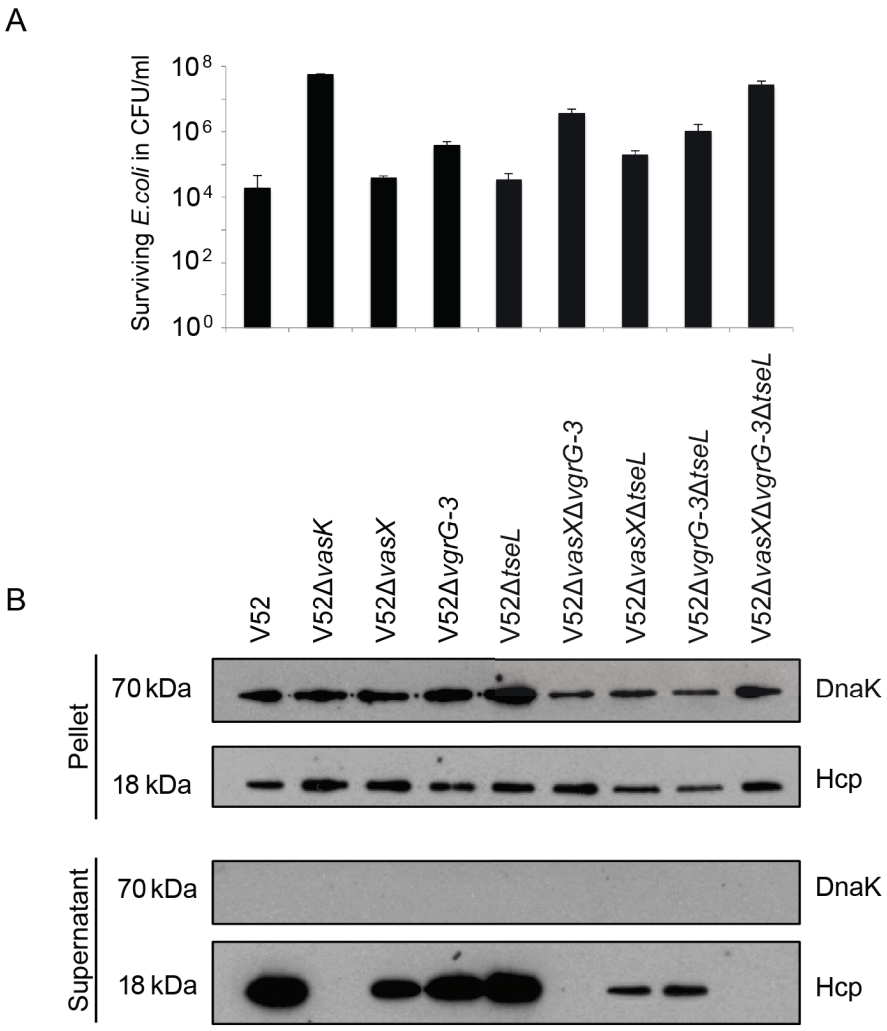


Figure 2-6 TseL, VasX, and VgrG-3 are required for killing of *E. coli* and Hcp secretion.

(A) Survival of rifampicin-resistant *E. coli* was determined by enumerating CFU following exposure to the indicated rifampicin-sensitive predator (listed on the x-axis). These data represent two independent experiments performed in technical duplicate. Error bars indicate the standard deviation. (B) V52Δ*vgrG-3*Δ*tseL* secretes Hcp. Pellet and supernatant samples were prepared using mid-logarithmic cultures of the strains indicated at the top of the blot. Samples were subjected to SDS-PAGE followed by western blotting with Hcp and DnaK (loading and lysis control, respectively) antibodies. Molecular weight is noted to the left of the blot. These data represent three independent experiments.

As expected, wild-type V52 expressed and secreted Hcp, whereas the *vasK* deletion mutant expressed but did not secrete Hcp (Figure 2-6). V52 with single deletions of *vasX*, *vgrG-3*, or *tseL* secreted Hcp. The lack of *vasX* or *vgrG-3* reduced Hcp secretion (Figure 2-6, Figure 9-1). Reduced levels of Hcp secretion were detected in the V52 double mutants that lack *vgrG-3* or *vasX* in addition to *tseL*. In the supernatant sample of mutants lacking *vasX* and *vgrG-3* or all three effector-encoding genes, no Hcp was detected. One possible explanation for this phenotype could be a mechanism that guarantees effector loading prior to ejection. DnaK was not detected in any supernatant sample, indicating that the samples were prepared properly and Hcp detected in the supernatant was the result of an active T6SS and not of cell lysis.

The results of the western blot analysis to monitor Hcp secretion suggest that *tseL*, *vasX*, and *vgrG-3* are required for a fully functional T6SS apparatus. This needs to be considered when interpreting the role of these genes in T6SS-mediated killing of *E. coli*. Because the effects of the gene deletions on Hcp secretion differ from their effects on *E. coli* killing, it seems likely that the proteins encoding the genes *tseL*, *vasX*, and *vgrG-3* function as antiprokaryotic effectors, beyond being required for the structural integrity of the T6SS. Mutants with catalytically inactive alleles of the effector protein-encoding genes could be used to test the effector activity independent of the role of the effector for the structural integrity of the T6SS.

2.3 Discussion

The three genes identified in this work, *tsiV1, tsiV2*, and *tsiV3* (Figure 2-7), are required for *V. cholerae* strain C6706 to be immune to a T6SS-mediated attack of the *V. cholerae* strain V52. Each gene depends on the nucleotide sequence of the gene encoded upstream to mediate protection. Further analysis revealed that TsiV1, TsiV2, and TsiV3 protect against TseL, VasX, and VgrG-3 proteins, respectively. No cross-immunity was observed. *V. cholerae* V52 genes *tseL*, *vasX*, and *vgrG-3* were shown to be necessary for *E. coli* killing and for the structural integrity of the V52 T6SS.

Based on the genetic analysis, my results suggest that TseL, VasX, and VgrG-3 are T6SS effector proteins required for bacteria-bacteria killing. This work complements

the biochemical analysis of these proteins and their antiprokaryotic activity performed by Sarah Miyata and Teresa Brooks in our laboratory and by other groups (Brooks et al., 2013; Dong et al., 2013; Miyata et al., 2013; Russell et al., 2013). Sarah Miyata revealed that VasX disrupts the inner membrane potential of target bacteria and proposed a model in which VasX acts as a pore-forming effector (Miyata et al., 2013). Teresa Brooks showed in a zymogram that VgrG-3 has peptidoglycan-degrading activity (Brooks et al., 2013). T6SS effectors with peptidoglycan-degrading activities are also found in other bacterial species in which they were characterized in depth as amidases, glycosidases, and muramidases (Russell et al., 2011; Russell et al., 2014a). TseL belongs to a group of T6SS effectors with lipase activity (Dong et al., 2013; Russell et al., 2013). In summary, TseL, VasX, and VgrG-3 target structures in the periplasm of target bacteria which leads to cell lysis. Some effectors like VasX also have been described to have antieukaryotic activity (Miyata et al., 2011). Lipase effectors of *P. aeruginosa* were shown to activate the Akt/PI3K pathway in HeLa cells (Jiang et al., 2014). Possible antieukaryotic activity of TseL requires further attention.

This variety of T6SS effectors is specifically inhibited by immunity proteins, which I identified in a transposon screen. How immunity proteins protect from T6SS effectors was further tested by Teresa Brooks, who showed that TsiV3 directly binds VgrG-3 using size exclusion chromatography. Furthermore, T. Brooks showed that TsiV3 inhibits the peptidoglycan-degrading activity of VgrG-3 by measuring the percent cell lysis of bacteria with a permeabilized outer membrane treated with or without purified TsiV3 in addition to purified VgrG-3 (Brooks et al., 2013). These findings were further supported by Zhang and colleagues who cocrystallized the C-terminal end of VgrG-3 in complex with a TsiV3 dimer and identified single amino acids within TsiV3 required for the dimerization and subsequently for protection from VgrG-3 (Zhang et al., 2014). Altogether, immunity proteins directly bind T6SS effectors to inhibit their toxic function.

Independently, Tao Dong also identified the genes with the locus tag VC1419, VCA0021, and VCA0124 that encode immunity proteins and introduced the nomenclature *tsiV1* (T6SS immunity gene in *V. cholerae* 1), *tsiV2*, and *tsiV3*, respectively, that we have adopted (Dong et al., 2013). Tao Dong identified *tsiV1-3* by

transposon mutagenesis and subsequent deep sequencing of the transposon mutants (Dong et al., 2013). Therefore, he prepared transposon mutant libraries of the strain V52 with a functional T6SS and of the strain V52Δ*hcp-1*Δ*hcp-2* with a dysfunctional T6SS. After growing the transposon mutants overnight at 37 ºC, the number of surviving mutants with a transposon insertion in a respective gene was quantified by deep sequencing of the transposon insertion junctions. Mutants with transposon insertions in *tsiV1*, *tsiV2,* and *tsiV3* were detected only in V52Δ*hcp-1*Δ*hcp-2* with a dysfunctional T6SS but not in wild type V52 because a lack of immunity proteins TsiV1, TsiV2 or TsiV3 renders bacteria susceptible to T6SS-mediated killing by neighbouring bacteria that have a functional T6SS. Tao Dong's results confirm our results and validate our experimental approach of, first, screening for C6706 transposon mutants susceptible to a T6SS-mediated attack of bacteria from the strain V52, and second, identifying the effector protein-encoding genes nearby. However, there are a number of limitations to our approach for identifying T6SS effector and immunity proteins, three of which I discuss in detail. First, immunity proteins encoded in the genome outside of the T6SS gene cluster would have been missed in our candidate approach that only focused on genes withinin the T6SS clusters. Second, we identified TsiV1, TsiV2 and TsiV3 because their cognate effectors TseL, VasX and VgrG-3, respectively, are toxic to C6706. T6SS effectors that are not toxic to C6706, for example because C6706 lacks the substrate for the effector, but are toxic to bacteria of other species that have the substrate would have not been detected in our screen. Third, the dual function of T6SS proteins as effectors and structural components complicates the validation of effector proteins in killing assays. For example, reduced killing cannot be clearly attributed to either lost T6SS effector function or to an overall decrease in T6SS formation. For these reasons, we missed the antiprokaryotic T6SS effector encoded by *tseH* (VCA0285) (Altindis et al., 2015). *TseH* is located outside of the T6SS gene clusters. Altindis and colleagues identified TseH in the culture supernatant of the *V. cholerae* strain 2740-80 but not in the supernatant of a mutant without a functional T6SS. TseH is predicted to have peptidoglycan-degrading activity and is required for killing of *E. coli* (Altindis et al., 2014; Shneider et al., 2013). In addition to TseH, the effectors TseL, VasX and VgrG-3 were detected in the culture supernatant of *V. cholerae* 2740-80 (Altindis et al., 2015).

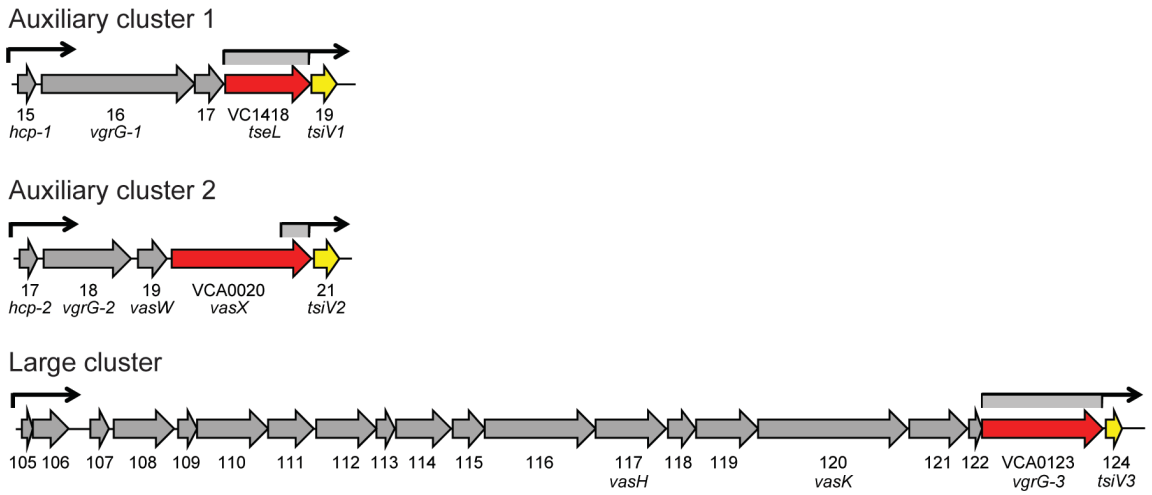


Figure 2-7 Dual regulatory profile of T6SS immunity protein-encoding genes.

Schematic representation of the *V. cholerae* T6SS gene clusters. Toxin-encoding genes are shown in red and immunity protein-encoding genes are shown in yellow. Promoters upstream of the individual T6SS clusters are indicated by arrows. Regions with promoter activity that drive expression of immunity protein-encoding genes are indicated in grey below arrows.

An additional result from our transposon screen was the dual regulation of T6SS immunity protein-encoding genes. Sarah Miyata further defined a region within *vasX* that contains an internal promoter region driving the expression of *tsiV2* under conditions in which the genes encoding structural and effector proteins of the T6SS are not expressed (Miyata et al., 2013). In addition to this internal promoter, the promoter upstream of *hcp* drives expression of all T6SS genes in the auxiliary cluster, including *tsiV2* (Bernard et al., 2011). The ability to not employ a functional T6SS and to maintain the expression of immunity proteins has at least three advantages. First, T6SS components are immunogenic (Ma and Mekalanos, 2010). To avoid secretion of T6SS components in situations where it is not needed may provide a strategy to hide from the immune system during infection. Second, bacteria like *P. aeruginosa* react to a T6SS-mediated attack of *V. cholerae* with a counterattack that might end up deadly for the *V. cholerae* bacterium itself (Basler et al., 2013). To avoid disturbing other bacteria can protect a microbe from a microbial counterattack. Third, the dual regulation of the T6SS enables *V. cholerae* to live in a mixed population of *V. cholerae* bacteria of which some but not all engage in

T6SS-mediated killing. The T6SS activity of *V. cholerae* has been described as "nondirectional" (Basler et al., 2013). Expression of immunity proteins could protect a subpopulation of *V. cholerae* bacteria that might save energy by not having an active T6SS. Having a subpopulation with an active T6SS could potentially protect the whole population. These scenarios indicate how the dual regulation of T6SS proteins can be advantageous. Differences in expression levels between individual T6SS immunity protein-encoding genes have been observed in *V. cholerae* colonizing the small intestine of infant rabbits (Fu et al., 2013). Even though C6706 represses its T6SS under laboratory conditions (MacIntyre et al., 2010; Zheng et al., 2010), the T6SS immunity protein-encoding genes are expressed in bacteria of the strain C6706 under laboratory conditions. It would be interesting to explore the regulatory network of C6706 T6SS immunity protein-encoding genes to find conditions under which they are not expressed. Dual regulation of immunity-protein encoding genes is not universally found among bacteria with a T6SS. In *Salmonella enterica* serovar Typhimurium, the T6SS effector and immunity protein encoding genes are both repressed under laboratory conditions by histone-like nucleoid structuring (H-NS) proteins (Brunet et al., 2015).

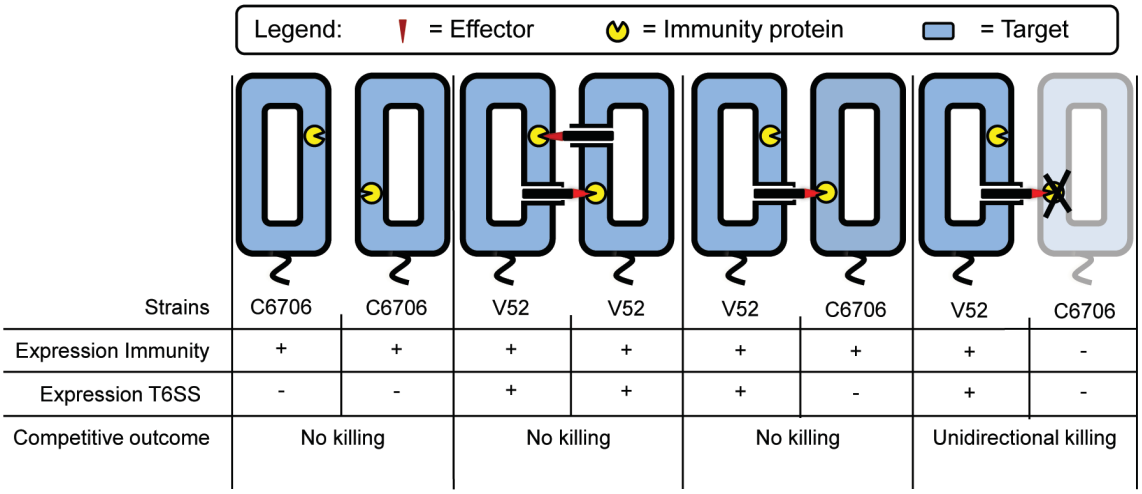


Figure 2-8 T6SS-mediated competition between bacteria of the strains V52 and C6706.

Two *V. cholerae* strains, C6706 and V52, exemplify the dual regulatory profile of the T6SS system. In the T6SS-off state, C6706 expresses only the T6SS immunity proteins (first panel). In contrast, V52 expresses structural, effector, and immunity T6SS proteins, the latter of which provide protection from T6SS-active kin bacteria (second panel). When T6SS-active V52 come in contact with C6706 in the T6SS-off state, C6706 is protected from a T6SS-mediated attack without engaging in T6SS-mediated virulence (third panel). C6706 succumbs to killing by V52 when immunity genes are removed or not expressed (fourth panel).

These findings advance our understanding of the mechanistic basis of T6SS-mediated antiprokaryotic killing (Figure 2-8). *V. cholerae* of the strain C6706 express immunity proteins but do not employ an active T6SS under laboratory conditions, thus C6706 bacteria peacefully coexist (Panel 1). *V. cholerae* of the strain V52 translocate T6SS effector proteins into each other and these effector proteins are inhibited by immunity proteins and allow the bacteria to also peacefully coexist (Panel 2). The expression of immunity proteins protects C6706 when exposed to bacteria of the strain V52 with an active T6SS (Panel 3). Deletion of C6706 immunity proteins results in T6SS-mediated unidirectional killing of C6706 by bacteria of the strain V52 (Panel 4).

Chapter 3

V. cholerae employs its T6SS for intraspecific competition among environmental isolates

Parts of this chapter have been published as:

Unterweger, D.*, M. Kitaoka*, S.T. Miyata, V. Bachmann, T.M. Brooks, J. Moloney, O. Sosa, D. Silva, J. Duran-Gonzalez, D. Provenzano and S. Pukatzki, Constitutive type VI secretion system expression gives *Vibrio cholerae* intra- and interspecific competitive advantages. PLoS One, 2012. 7(10):e48320.

*These authors contributed equally to this publication

Daniele Provenzano provided Figure 3-1, Maya Kitaoka provided Figure 3-3, and Sarah Miyata provided Figure 3-4.

3 *V. cholerae* employ a T6SS for intraspecific competition among environmental isolates

3.1 Introduction

V. cholerae is abundant in the aquatic environment. In addition to multiple *V. cholerae* strains, the aquatic environment harbors a wide variety of related bacterial species such as *Vibrio harveyi* (Austin and Zhang, 2006), *Vibrio communis* (Chimetto et al., 2011), and *Pseudoalteromonas phenolica sp.* (Isnansetyo and Kamei, 2003).

So far, the T6SS of *V. cholerae* has been mainly studied in clinical isolates of *V. cholerae*, like the O37 serogroup strain V52 and the O1 serogroup strain C6706. The T6SS gene cluster, however, is not only found in the genomes of clinical isolates but in all yet sequenced genomes of *V. cholerae* strains independent of their source of isolation (Boyer et al., 2009). *V. cholerae* strains SCE226 and SCE223 isolated from the environment in Calcutta (India) have a functional T6SS under laboratory conditions and were not yet tested for their ability to kill other bacteria (Faruque et al., 2003a; Pukatzki et al., 2006). We hypothesized that freshly isolated *V. cholerae* strains from the Rio Grande Delta employ a functional T6SS to kill other bacteria.

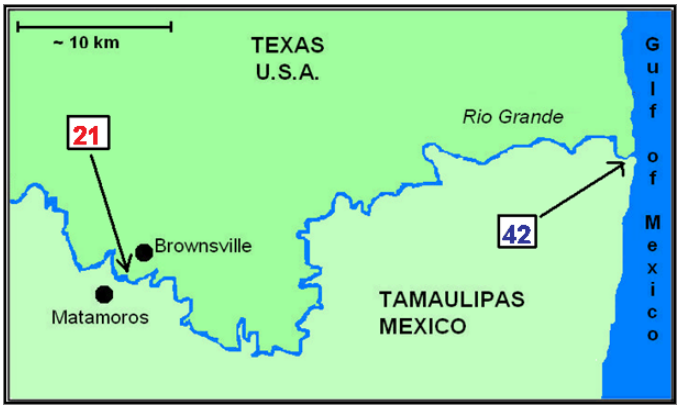


Figure 3-1 Map indicating the sample sites of Rio Grande *V. cholerae*.
Geographic map of Brownsville and the Rio Grande outfall into the Gulf of Mexico. Sample site 21 of isolates DL2111 and DL2112 close to the city Brownsville is indicated in red. Sample site of isolates DL4211 and DL4215 close to the river is indicated in blue.

The Rio Grande Delta at the border between Texas (USA) and Mexico contains environmental *V. cholerae* and is an accessible area to take fresh isolates. Daniele

Provenzano and colleagues took samples from two separate locations proximal and distal to the outfall of the river (Figure 3-1). These Rio Grande *Vibrio cholerae* (RGVC) isolates provide the basis for our following analysis in which we investigate the activity and antibacterial properties of their T6SS.

3.2 Results

3.2.1 RGVC isolates exhibit T6SS-mediated antimicrobial properties

The *V. cholerae* O37 serogroup strain V52 derived from a clinical isolate was previously shown to use its T6SS to kill *E. coli* and *Salmonella* Typhimurium (MacIntyre et al., 2010) among other Gram-negative species. To determine the role of the T6SS in microbial competition of environmental strains, we tested *V. cholerae* isolated from two different sites of the Rio Grande: isolates from the river close to the city of Brownsville, and isolates from the river outfall into the Gulf of Mexico (Figure 3-1). We noticed that the isolates taken close to the city lack the O-antigen found in the outer membrane of Gram-negative bacteria. Such strains are described as "rough" and contain an altered lipopolysaccharide (LPS) core oligosaccharide with no O-antigen attached (De et al., 2004). Such a condition was also observed among isolates from hospitalized patients with diarrhoea(Mitra et al., 2001) and is most likely due to pressure by phages that use LPS as their receptor. We assessed the T6SS of two separately isolated but genetically identical isolates DL2111 and DL2112 (as determined by deep sequencing [Illumina platform] of a polymorphic 22 kb fragment [Genbank accession numbers JX669612 and JX669613]) to minimize the chance of phenotypic variation due to genetic exchange.

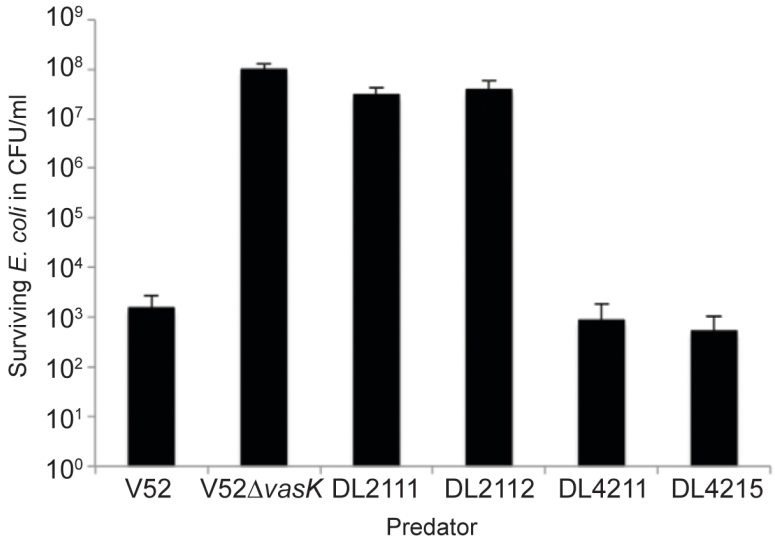


Figure 3-2 Ability of RGVC isolates to kill *E. coli*.

RGVC isolates were tested for their ability to engage in T6SS-mediated prokaryotic killing. V52 and V52Δ*vasK* were used as virulent and avirulent controls, respectively. V. cholerae and E. coli were mixed in a 10:1 ratio and incubated for 4 hours at 37 ºC. Bacterial spots were resuspended, serially diluted, and plated on *E. coli*-selective media to determine the number of surviving E. coli. The averages and standard deviations of two independent experiments, each performed in duplicate, are shown.

To determine whether environmental RGVC are capable of killing bacteria, I performed an *E. coli* killing assay (Figure 3-2). RGVC isolates and the *E. coli* strain MG1655 were spotted on LB nutrient agar plates, and the number of surviving MG1655 cells was determined after a 4 hour incubation at 37 °C. V52 and V52Δ*vasK* were used as virulent and avirulent controls, respectively. The presence of V52 resulted in an average ~ 5-log reduction of viable *E. coli*. The isolates DL4211 and DL4215 killed *E. coli* at levels comparable to V52 (Figure 3-2). In contrast, both isolates DL2111 and DL2112 were unable to kill *E. coli* prey. In summary, RGVC isolated at the river outfall readily killed *E. coli* while RGVC isolated next to Brownsville appeared to be attenuated.

3.2.2 Expression of Hcp in RGVC isolates

Next, we set out to test whether RGVC isolates were able to produce and secrete the T6SS hallmark protein Hcp, because experimental results presented thus far suggested that *V. cholerae*'s ability to kill bacterial competitors could be mediated by the T6SS. As shown in Figure 3-3, isolates DL4211 and DL4215 produced Hcp at sufficient levels to be detected by western blots probed with Hcp antiserum. In contrast, isolates DL2111 and DL2112 did not produce or secrete Hcp. The presence of Hcp correlated with virulence as the isolates DL4211 and DL4215 secreted Hcp (Figure 3-3) and killed *E. coli* (Figure 3-2), while isolates DL2111 and DL2112 did not produce Hcp and did not kill *E. coli*. Because DL2111 and DL2112 lack an active T6SS, we used the strains with an active T6SS, DL4211 and DL4215, in subsequent experiments.

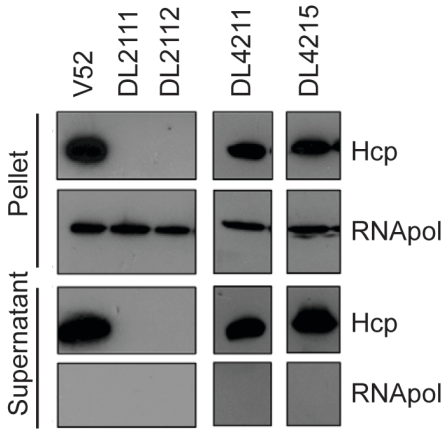


Figure 3-3 RGVC isolates differ in T6SS regulation.

Indicated RGVC isolates and V52 (positive control) were cultured to midlogarithmic phase of growth followed by centrifugal separation of pellets and culture supernatants. Supernatant portions were concentrated by TCA precipitation and both fractions were subjected to SDS-PAGE followed by western blotting using the antibodies indicated. Experiments were repeated at least three times with similar results.

3.2.3 T6SS-active RGVC isolates use their T6SS to compete with environmental isolates of other species and genera

Because RGVC isolates with active T6SSs kill *E. coli*, we hypothesized that RGVC isolates use their T6SS to compete with other bacteria in their environmental niche. To test this hypothesis, we isolated three environmental bacterial non-*V. cholerae* strains from estuaries where the Rio Grande meets the Gulf of Mexico. Sequencing of 16S-rRNA identified these bacterial species as *Vibrio communis*, *Vibrio harveyi*, and *Pseudoalteromonas phenolica*. We then tested whether DL4211 and DL4215 were able to kill these environmental bacteria in a T6SS-dependent fashion. As shown in Figure 3-4, both DL4211 and DL4215 killed all three environmental isolates. The observed killing required a functional T6SS, as *vasK* deletion mutants lost their ability to kill. Killing of the environmental bacteria was restored by complementing the *vasK* mutant backgrounds with episomally expressed *vasK in trans*. Therefore, we propose that constitutive expression of T6SS genes provides some RGVC isolates with the means to kill bacteria of species and genera other than *Vibrio.*

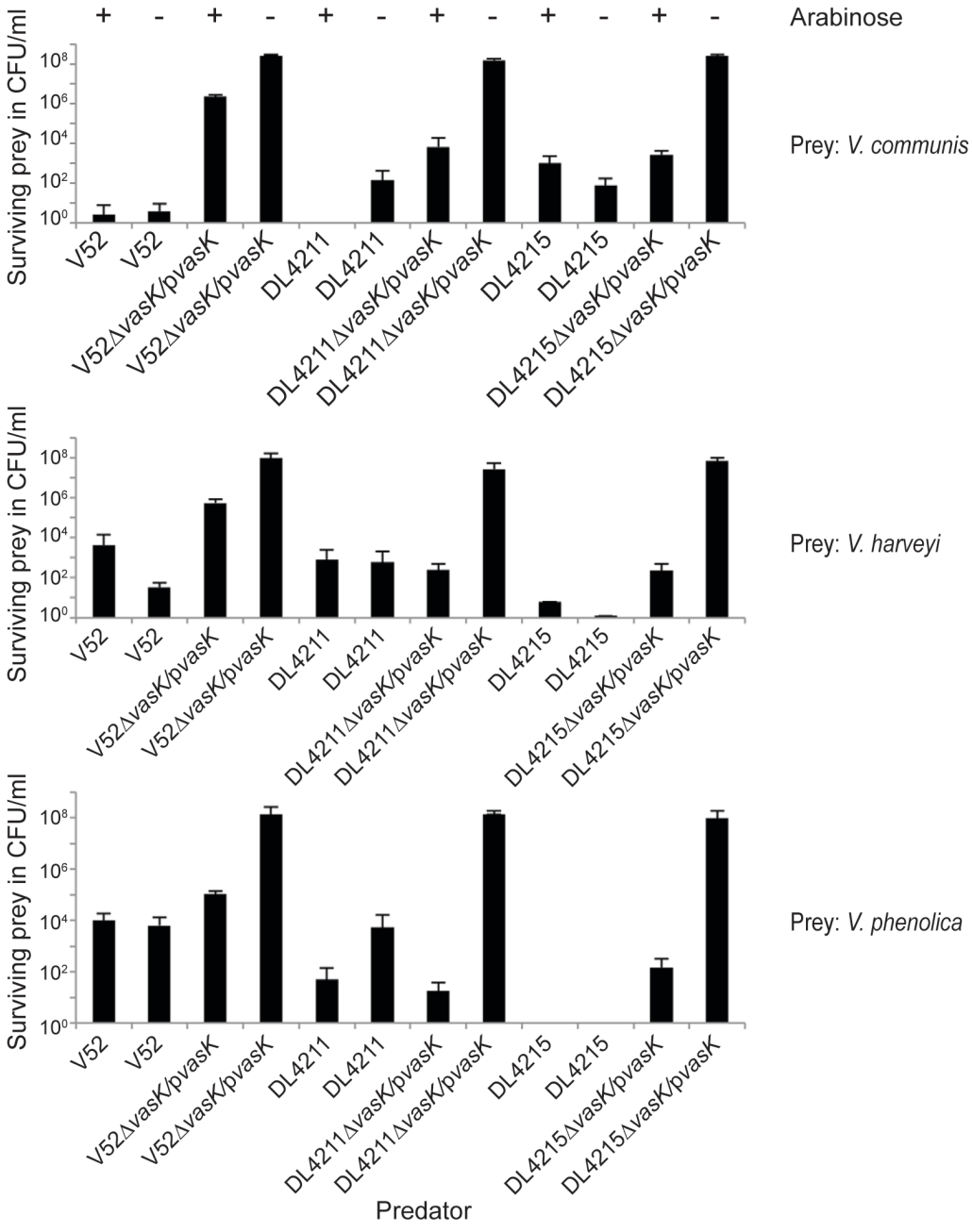


Figure 3-4 RGVC isolates kill bacterial neighbours.

V. cholerae and prey bacteria were mixed at a 10:1 ratio and incubated on ½ YTSS agar for 4 hours at 30 ºC. Bacterial spots were resuspended, serially diluted, and plated on selective YTSS agar to determine the number of surviving prey. The average and standard deviations of three independent experiments, each performed in duplicate, are shown.

3.2.4 RGVC isolates with an active T6SS engage in intraspecific competition

As RGVCs killed close relatives such as *V. harveyi* (Figure 3-4), we wondered if the RGVC isolates have the ability to kill each other. To test this hypothesis, we mixed V52, DL4211, and DL4215 (predators) with four different RGVC isolates as prey bacteria. To eliminate the killing activity of T6SS+ prey, we used *vasK*-deficient mutants with a disabled T6SS as prey. DL2111 and DL2112 wild-type RGVC isolates were used as prey since they do not express Hcp (Figure 3-3) and are thus T6SS-negative. Following a 4-hour coincubation, we determined the number of surviving prey. T6SS-negative prey bacteria were not killed by their isogenic T6SS+ parent strain, but were killed by other T6SS+ isolates (Figure 3-5). Exposure to a predator with a disabled T6SS resulted in ~ 10^8 surviving prey bacteria. Similar numbers of surviving prey were obtained when the prey were mixed with an isogenic strain that was marked with a different antibiotic resistance cassette. Thus, killing of T6SS-negative prey required a functional T6SS. Surprisingly, the *vasK* mutant of DL4215 displayed virulence toward V52Δ*vasK*, but not against DL4211Δ*vasK* or a differently marked DL4215Δ*vasK* kin strain (Figure 3-5). Since DL4215Δ*vasK* does not kill *V. communis, V. harveyi*, or *P. phenolica* (Figure 3-4), we hypothesize that DL4215 exhibits selective T6SS-independent antimicrobial activity against V52Δ*vasK* (and probably V52).

In conclusion, *V. cholerae* uses its T6SS not solely for competition with bacteria of other species and genera (Figure 3-4) but also for competition within its own species (Figure 3-5).

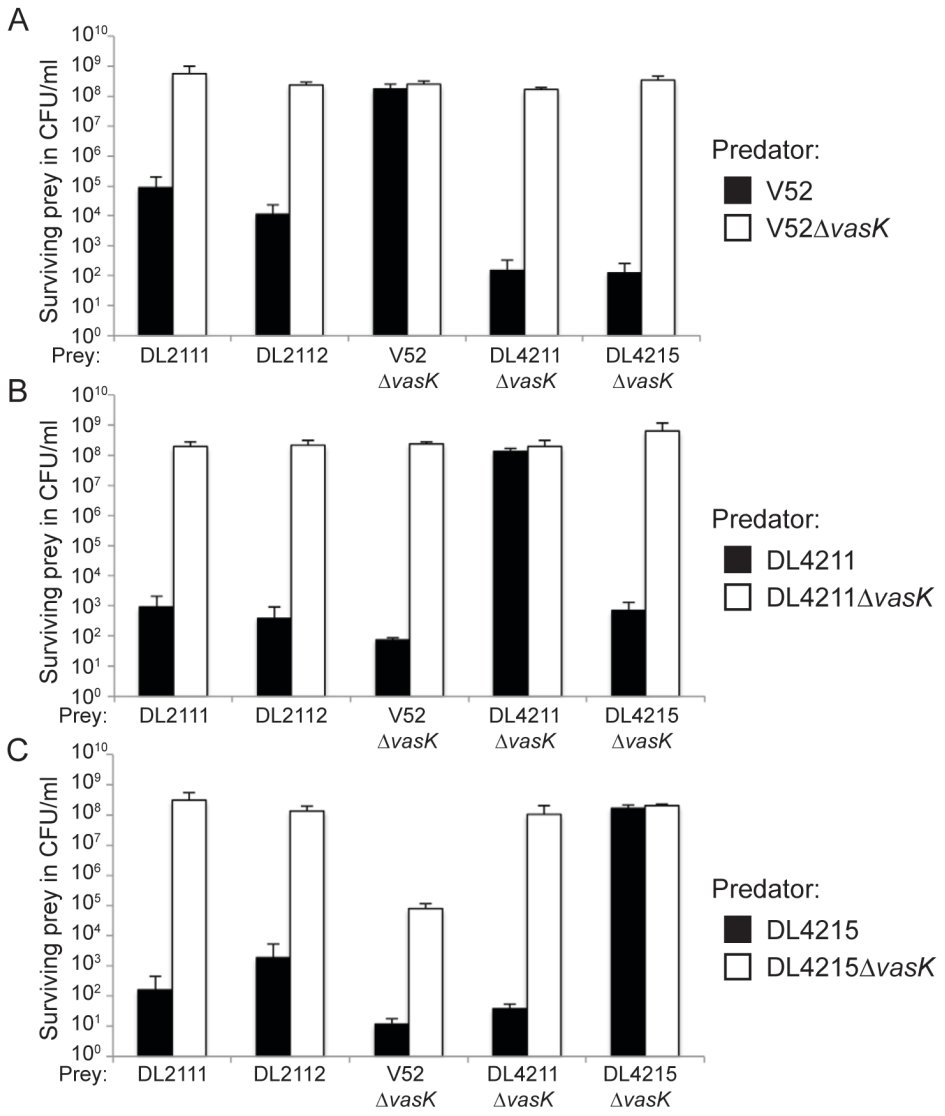


Figure 3-5 T6SS-dependent competition among *V. cholerae* isolates.

(A–C) T6SS-active *V. cholerae* isolates successfully competed with each other and outcompeted the isolates DL2111 and DL2112 in a T6SS-dependent manner. All combinations among the isolates and their isogenic *vasK* mutants were tested in a killing assay: Predator- and prey-*V. cholerae* were mixed in a 10:1 ratio and incubated for 4 hours at 37 ºC. Bacterial spots were resuspended, serially diluted, and plated on selective media to determine the number of surviving prey. The number of surviving prey in the presence of T6SS+ or T6SS- predator are shown.

3.3 Discussion

We used environmental isolates to characterize the *V. cholerae* T6SS activity against bacteria of the same and other species *V. cholerae* may encounter in environmental reservoirs (Figure 3-6). We observed that some but not all of the environmental *V. cholerae* isolates have an active T6SS under laboratory conditions. These observations support previous findings on the T6SS activity of environmental isolates from India (Pukatzki et al., 2006). *V. cholerae* isolates with an active T6SS were shown to engage in T6SS-mediated bacterial killing of *E. coli* and environmental isolates of other *Vibrio* species. Unlike previously tested *V. cholerae* strains, the environmental *V. cholerae* isolates were not immune to a T6SS-mediated attack of bacteria of the strain V52, and I demonstrated that they engage in T6SS-mediated intraspecific competition.

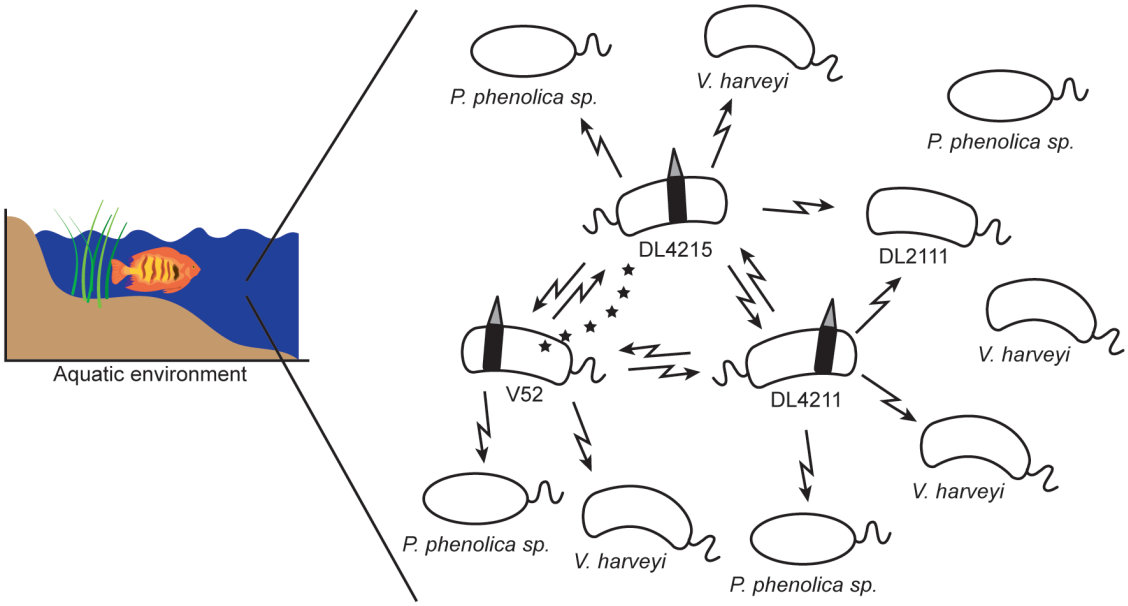


Figure 3-6 T6SS-mediated microbial competition of environmental *V. cholerae* isolates. Strain names for *V. cholerae* and species names for other bacteria are indicated below the graphical depiction of bacteria. Arrows between cells indicate T6SS-mediated killing in the direction of the killing. Stars indicate a putative T6SS-independent killing mechanism.

The observation that some but not all of the environmental *V. cholerae* isolates have an active T6SS under laboratory conditions further exemplifies the strain-specific

differences in the regulation of the T6SS gene cluster that is encoded in the genomes of all strains sequenced to date (Boyer et al., 2009). Maya Kitaoka further characterized the same *V. cholerae* isolates with an inactive T6SS and found a frameshift mutation at position 157 of *vasH* (Unterweger et al., 2012). VasH is encoded in the large T6SS gene cluster and required for T6SS gene expression by recruiting sigma factor 54. M. Kitaoka tried to activate the T6SS of the isolates DL2111 and DL2112 by expressing *vasH* from bacteria of the strain V52 *in trans*. Subsequently, she could restore Hcp expression but not secretion, indicating that further experiments on the regulation of the T6SS in *V. cholerae* strains with an active T6SS need to be conducted to determine the cause(s) of the lost T6SS gene expression and activity in isolates like DL2111 and DL2112 (Unterweger et al., 2012).

In addition to the differences in T6SS activity between isolates DL2111, DL2112 and DL4211, DL4215, the latter two isolates differ in their site of isolation and in their O-antigen. We can speculate about links between these phenotypes. *V. cholerae*, like DL2111 and DL2112 that lack their O-antigen, are less immunogenic during infection (Miller et al., 1972) and thus might not require an active T6SS to fight off the immune system of the host. The influence of humans on the environment close to the city of Brownsville might stress *V. cholerae* to an extent to which they cannot afford an active T6SS. Alternatively, compared to the river outfall, the urban environment might harbor a lower bacterial diversity in which there is no need for interbacterial competition. These speculations require large sample sizes and systematic analyses to be further tested.

The finding of T6SS-mediated killing between environmental *Vibrio* isolates of the same and other species isolated in the same environmental region strengthens the notion that antiprokaryotic activity is an important and physiologically relevant function of the T6SS. Further studies are needed to analyze T6SS-mediated killing and its consequences for the survival of *V. cholerae* in the environment. To test if T6SS-mediated competition determines the colonization of a niche, copepods isolated from the environment that serve as a natural reservoir for *V. cholerae* (Rawlings et al., 2007; Tamplin et al., 1990) could be analyzed for the *Vibrio* species they are colonized with. Multiple isolates from a single copepod and from multiple copepods could be mixed in killing assays and tested for T6SS-mediated competition. An observation of a single

copepod being only colonized by isolates that peacefully coexist whereas isolates from different copepods outcompeting each other would support a role of the T6SS in colonizing a niche.

When DL4215Δ*vasK* and V52Δ*vasK* were mixed, the latter were killed (Figure 3-5) despite the dysfunctional T6SS in DL4215Δ*vasK*. This finding suggests a T6SS-independent bacterial killing mechanism possibly mediated by a bacteriocin of *V. cholerae* for which examples have been described as "Vibriocins" (Israil et al., 1987). The presence of a bacterial killing mechanism other than the T6SS raises basic questions about the need and function of the T6SS in contrast to T6SS-independent bacterial killing mechanisms. To further characterize the T6SS-independent killing mechanism of DL4215, I propose experimental approaches to answer two key questions. First, what is the molecular mechanism of T6SS-independent killing by DL4215? The creation of a transposon library of the strain DL4215Δ*vasK* and a screen for a transposon insertion that loses the ability to kill V52Δ*vasK* could lead to the identification of a toxic effector protein. One limitation to this approach would be the redundancy of mutants. Alternatively, culture supernatants could be tested for antibacterial activity and a subsequent mass spectrometry analysis of the active fraction of the supernatant could reveal the identity of a toxic effector protein. Once the toxin(s) are identified, I suggest experiments to compare and contrast features of bacterial killing like target specificity, or contact-dependence that is required for T6SS-mediated killing and still needs to be characterized for the alternative killing mechanism. Second, how widespread is T6SS-independent killing among *V. cholerae* strains? Given that the genes and herein encoded proteins required for T6SS-independent killing are identified, these sequences could be searched for in all sequenced *V. cholerae* strains in a bioinformatics approach. Such an analysis would provide information on the importance of such a mechanism for the whole species or rather for isolates from a certain location. In summary, DL4215 employs a T6SS-independent bacterial killing mechanism that is worth further investigation. Elucidation of the DL4215 bacteria killing mechanism will help us to learn more about the competitiveness of *V. cholerae* and the individual molecular mechanisms it uses to kill other bacteria.

Analysis of environmental *V. cholerae* isolates from the Rio Grande Delta revealed T6SS-mediated intraspecific competition between different isolates. This observation raises the question of how intraspecific competition affects the species *V. cholerae.* Considering that we have identified immunity proteins in C6706 that protect this strain from an attack of V52, we wonder how intraspecific competition is achieved mechanistically. These questions form the basis for chapter 4 of this thesis.

Chapter 4

T6SS effector modules determine compatibility among *V. cholerae* strains

Portions of this chapter have been published as

Unterweger, D., Miyata, S.T., Bachmann, V., Mullins, T., Brooks, T., Kostiuk, B., Provenzano D. and Pukatzki S. , The *Vibrio cholerae* type VI secretion system employs diverse effector modules for intraspecific competition. Nature Communications, 2014. 5:3549.

Verena Bachmann performed killing assays shown in Figure 4-5A. Ashley Wilton performed the killing assays shown in Figure 4-5C. Teresa Brooks performed the western-blot shown in Figure 4-6A. Ben Kostiuk performed killing assays that are part of Figure 4-6D. Stefan Pukatzki created the phylogenetic tree shown in Figure 4-7. Travis Mullins performed the original anlaysis on CTX and TCP shown in Figure 4-9.

4 T6SS effector modules encode a diversity of *V. cholerae* T6SS effectors

4.1 Introduction

Over more than 100 years, *V. cholerae* bacteria have been isolated from patients and environmental reservoirs worldwide. The diversity of over 200 different serogroups among these isolates exemplifies the diversity of the species *V. cholerae*.

Only O1 serogroup strains that encode the virulence factors cholera toxin (CT) and toxin-coregulated pilus (TCP) have caused pandemic spread of the diarrhoeal disease cholera (Herrington et al., 1988; Holmgren, 1981; Taylor et al., 1987). TCP biosynthesis genes are encoded within *Vibrio* pathogenicity island-1 (VPI-1). The claim that VPI-1 is part of a prophage (Karaolis et al., 1999) could not be confirmed (Faruque et al., 2003c). TCP protruding from the bacterium serves as a receptor for the filamentous bacteriophage CTX-ϕ that transduces CT genes into the chromosomes of *V. cholerae* strains (Waldor and Mekalanos, 1996). The contribution of CTX-ϕ is crucial to pathogenesis as only *V. cholerae* strains encoding CT are capable of pandemic spread (Waldor and Mekalanos, 1996). When humans consume contaminated water, toxigenic *V. cholerae* passes through the gastric acid barrier and colonizes the small intestine (reviewed in (Almagro-Moreno et al., 2015)). Upon mucus penetration, *V. cholerae* attaches to the epithelium (Krebs and Taylor, 2011). Increasing concentrations of bicarbonate enhance the affinity of ToxT to DNA (Abuaita and Withey, 2009). ToxT activates Tcp and CT, which induces watery diarrhoea (DiRita et al., 1991). Next, *V. cholerae* multiplies rapidly and exits the human host during diarrhoeal purges, (Sack et al., 2004) returning *V. cholerae* to the environment. Schild and colleagues used a modified RIVET screen to identify genes that are expressed during late stages of infection (Schild et al., 2007). When testing the role of individual genes, which were expressed late, in an infection model, only a few mutants showed a colonization defect whereas most of the other identified genes contribute to bacterial survival in stool and pond water (Schild et al., 2007). Sequence tag-based analysis of microbial populations (STAMP) was recently performed by Abel and colleagues on *V. cholerae* infecting the infant rabbit (Abel et al., 2015). Differences in the size of the founder population at different sites of the intestine and at different time-points during infection indicate multiple bottlenecks.

When analyzing *V. cholerae* strains from the Rio Grande river delta, we first noticed that *V. cholerae* strains compete against each other in a T6SS-dependent manner (Unterweger et al., 2012). Because T6SS-mediated killing happens when two bacteria are in direct contact with each other (MacIntyre et al., 2010), T6SS-mediated killing between bacteria of two strains prevents their coexistence in direct contact.We subsequently defined two strains that kill each other on contact as "incompatible". In contrast, "compatible" strains are able to co-exist with each other in direct contact. One example is the compatible *V. cholerae* strains C6706 and V52 (MacIntyre et al., 2010). Bacteria of the strain C6706 employ immunity proteins to protect themselves against a T6SS-mediated attack of bacteria from the strain V52 (Brooks et al., 2013; Dong et al., 2013; Miyata et al., 2013). Immunity proteins thus provide the molecular basis for compatibility whereas the molecular mechanism underlying incompatibility among *V. cholerae* strains remained unknown. We hypothesized that bacteria of incompatible *V. cholerae* strains employ immunity proteins that protect from an attack by bacteria of the same strain but fail to protect from the T6SS effectors of a *V. cholerae* strain that harbors different T6SS effectors. Additionally, we proposed that intraspecific competition is not restricted to *V. cholerae* strains isolated from the Rio Grande but rather a general phenomenon of *V. cholerae* with implications on the evolution of this species because T6SS-based (in)compatibility restricts physical contact and genetic exchange. To test our hypotheses, we performed a bioinformatics analysis on an increased sample size of *V. cholerae* strains.

4.2 Results

4.2.1 Conserved T6SS gene clusters harbour diverse effector modules

The genome of *V. cholerae* strain V52 contains three immunity protein-encoding genes *tsiV1*, *tsiV2* and *tsiV3* at the 5'end of the T6SS gene clusters auxiliary cluster 1, auxiliary cluster 2 and the large cluster, respectively (Brooks et al., 2013; Dong et al., 2013; Miyata et al., 2013). The T6SS effectors are encoded in the respective genes upstream of the immunity protein-encoding genes. We hypothesized that the genomes of compatible and incompatible *V. cholerae* strains do not contain the same immunity protein-encoding genes.

To test this hypothesis, the nucleotide sequences of the three T6SS clusters of 37 *V. cholerae* strains collected from a variety of sources over the past 77 years (Table of bacterial strains in Chapter 8) were aligned and compared. The separate gene cluster encoding *tseH* and *tsiH* are not included in this analysis because they were not yet identified to encode T6SS effector and immunity proteins when we conducted this analysis. An overview over the conservation of *tseH* and *tsiH* is shown in the appendix (Figure 9-4). As shown in Figure 4-1, auxiliary clusters 1 and 2, and the large cluster were found to be highly conserved among the available *V. cholerae* genomes except for regions of high diversity and low GC content (shaded blue) that contain genes encoding the bacterial effectors TseL, VasX, and VgrG-3, the adaptors of the cognate cargo effectors, and their immunity proteins TsiV1, TsiV2 and TsiV3, respectively.

Conserved T6SS genes contained an average GC content of ~ 51%, whereas the GC content in the region coding the effector immunity pairs was an average of ~ 9% lower (Figure 4-1) suggesting that these genes have a different origin than the remainder of the T6SS gene clusters and may have been acquired independently.

We named these regions of high diversity and low GC content 'effector modules', each encoding a T6SS effector and its cognate immunity protein. The combination of all three effector modules in one strain comprises an effector module set.

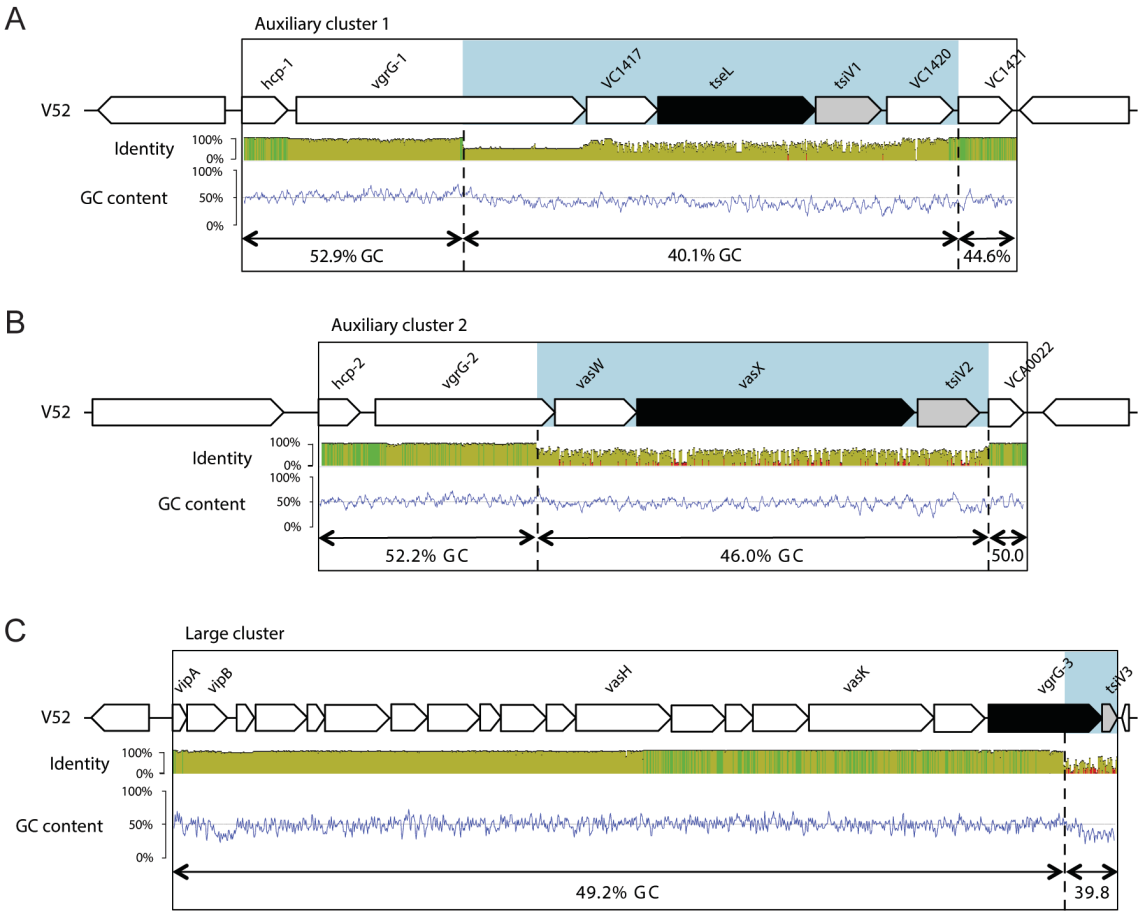


Figure 4-1 T6SS effector modules are highly diverse and differ in GC content

The boxes highlight auxiliary cluster 1 (a), auxiliary cluster 2 (b) and the large gene cluster (c) that encode the T6SS of *V. cholerae* V52. The results of a sliding-window analysis (with a window size of 10 nucleotides) in which the T6SS clusters of 37 *V. cholerae* strains were aligned and compared is indicated as average percent of nucleotide identity: 100% identity (green bars), at least 30% identity (yellow bars) and less than 30% identity (red bars). The GC content of *V. cholerae* V52 T6SS-encoding genes is plotted as the result of a sliding-window analysis (with a window size of 40). Average percentage of GC content was calculated for indicated regions. The T6SS effector genes for bacteria–bacteria interactions are shown in black and their cognate immunity genes are shown in grey. Regions of high diversity and low GC content are highlighted in blue. Reprinted by permission from McMillan Publishers Ltd: Nature Communications, Unterweger et al., 2014, copyright 2014.

4.2.2 *V. cholerae* strains harbour a diversity of putative T6SS effector and immunity proteins

The bioinformatics analysis among the T6SS gene clusters of multiple *V. cholerae* strains revealed a diversity of T6SS effector modules. The proteins encoded by genes located in these effector modules were characterized as effector and immunity proteins in the strain V52 (see chapter 1 and (Brooks et al., 2013)). The proteins encoded in genes in effector modules of other strains were previously annotated as uncharacterized, hypothetical proteins. We hypothesized that, like in V52, other strains encode T6SS effector and immunity proteins in their effector modules.

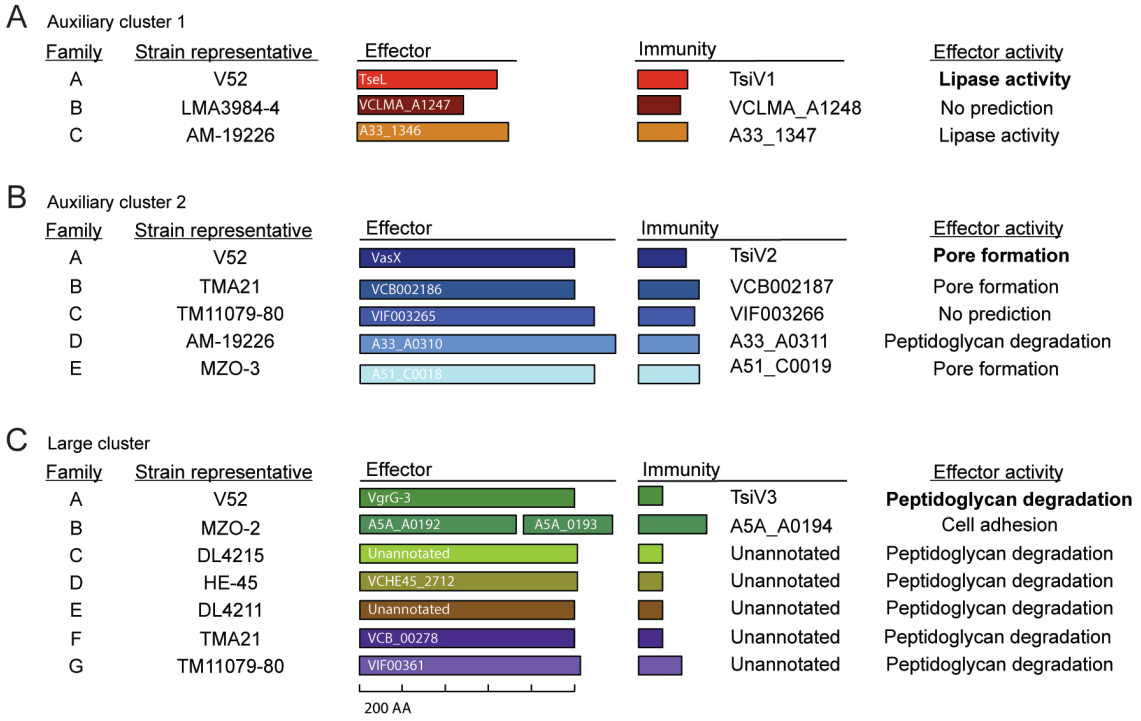


Figure 4-2 Families of *V. cholerae* T6SS effector modules

Effector and immunity proteins of individual families encoded in auxiliary cluster 1 (a), auxiliary cluster 2 (b) and the large cluster (c) are indicated by coloured boxes (box length is proportional to predicted protein length). The family of the effector module encoding the effector and immunity genes and a representative strain are shown on the left. Confirmed (bold type) or predicted (normal type) effector activities are shown on the right. Reprinted by permission from McMillan Publishers Ltd: Nature Communications, Unterweger et al., 2014, copyright 2014.

To identify putative T6SS effectors, I extracted the longer of the two open-reading frames in each effector module and translated the nucleotide sequence. If similar amino acid sequences were derived from multiple effector modules, one representative sequence was chosen per group. This sequence was then analyzed for conserved domains and homologous proteins with known function using the bioinformatics tool BLAST (Altschul et al., 1990). In addition, the amino acid sequence was analyzed by the online tool for protein modeling, prediction and analysis Phyre2 to identify proteins with structural similarities (Kelley et al., 2015). Figure 4-2 shows the results of our analysis. Genes of the effector modules encode proteins with predicted peptidoglycan degrading, pore-forming and lipase activity. In a few cases, no prediction could be made. In general, there seems to be a trend that genes found in effector modules in auxiliary cluster 1 encode lipase effectors, genes of effector modules in auxiliary cluster 2 encode pore forming effectors and genes of effector modules in the large cluster encode peptidoglycan degrading effectors. An exemplary exception to this rule is a gene in the effector module of auxiliary cluster 2 of the strain AM-19226 encoding a predicted peptidoglycan-degrading and not a predicted pore-forming effector.

The nucleotide sequences of short and sometimes unannotated open-reading frames adjacent to and sometimes overlapping with putative effector-protein encoding genes were also translated into the amino acid sequence. The resulting proteins are of unknown function but share a similar length than the amino acid sequence of immunity proteins with known function TsiV1, TsiV2 and TsiV3. Analyses of the diversity of amino acid sequences encoded in the short open reading frames showed that these proteins are very diverse and segregate in groups similarly to the putative effector proteins.

These results suggest that effector modules encode a variety of putative effector and immunity proteins.

4.2.3 Assigning compatibility groups based on T6SS immunity proteins

The diversity of effector modules and the therein encoded proteins caused the need for an intuitive way of indicating the effector module set (made up of three

modules) of a given *V. cholerae* strain. We decided to label an effector module based on the therein encoded immunity protein, because the ability of the immunity protein to protect from a given T6SS effector determines survival or death of the harboring bacterium when encountering a T6SS-mediated attack of another bacterium (see chapter 2, (Unterweger et al., 2014)). Immunity proteins are expected to differ from each other if they inhibit a different effector protein. One example for *V. cholerae* strains that breach immunity are O1 *V. cholerae* strains of the classical biotype.

To group effector modules based on similarities and differences between the immunity proteins, we analyzed the effector modules of one gene cluster starting with auxiliary cluster 1. Amino acid sequences of 37 TsiV1 immunity proteins encoded in the respective effector module in auxiliary cluster 1 were aligned. For each group of identical amino acid sequences, one representative sequence was kept for a second alignment of unique amino acid sequences (12/37). I then determined the homology (in percent identity) between the individual amino acid sequences and plotted the resulting values (Figure 4-3). The first column shows the percentage identity of all 12 unique amino acid sequences in respect to the amino acid sequence of the immunity protein encoded in a gene in the effector module in auxiliary cluster 1 of the strain MZO-2. As indicated by the blue dot, the amino acid sequence of MZO-2 is 100% identical to itself. The amino acid sequence of O395 differs from MZO-2 by two amino acid substitutions and the two sequences thus share 99.2% identity. The amino acid sequences of strains other than V52 and O395 only share about 10% sequence identity with the sequence of the O14 serogroup strain MZO-2 isolated from a cholera patient in Bangladesh in 2001. In addition to MZO-2, the same comparison was performed for the other 11 amino acid sequences. Because proteins sharing more than 30% amino acid identity are predicted to fold similarly (Rost, 1999), I then grouped sequences like the ones of MZO-2, V52 and O395 that share more than 30% sequence identity into one family, indicated by a box in Figure 4-3A. Different families were labeled with different alphabetical letters A, B and C. Sequences from different families share less than 30% amino acid identity in common.

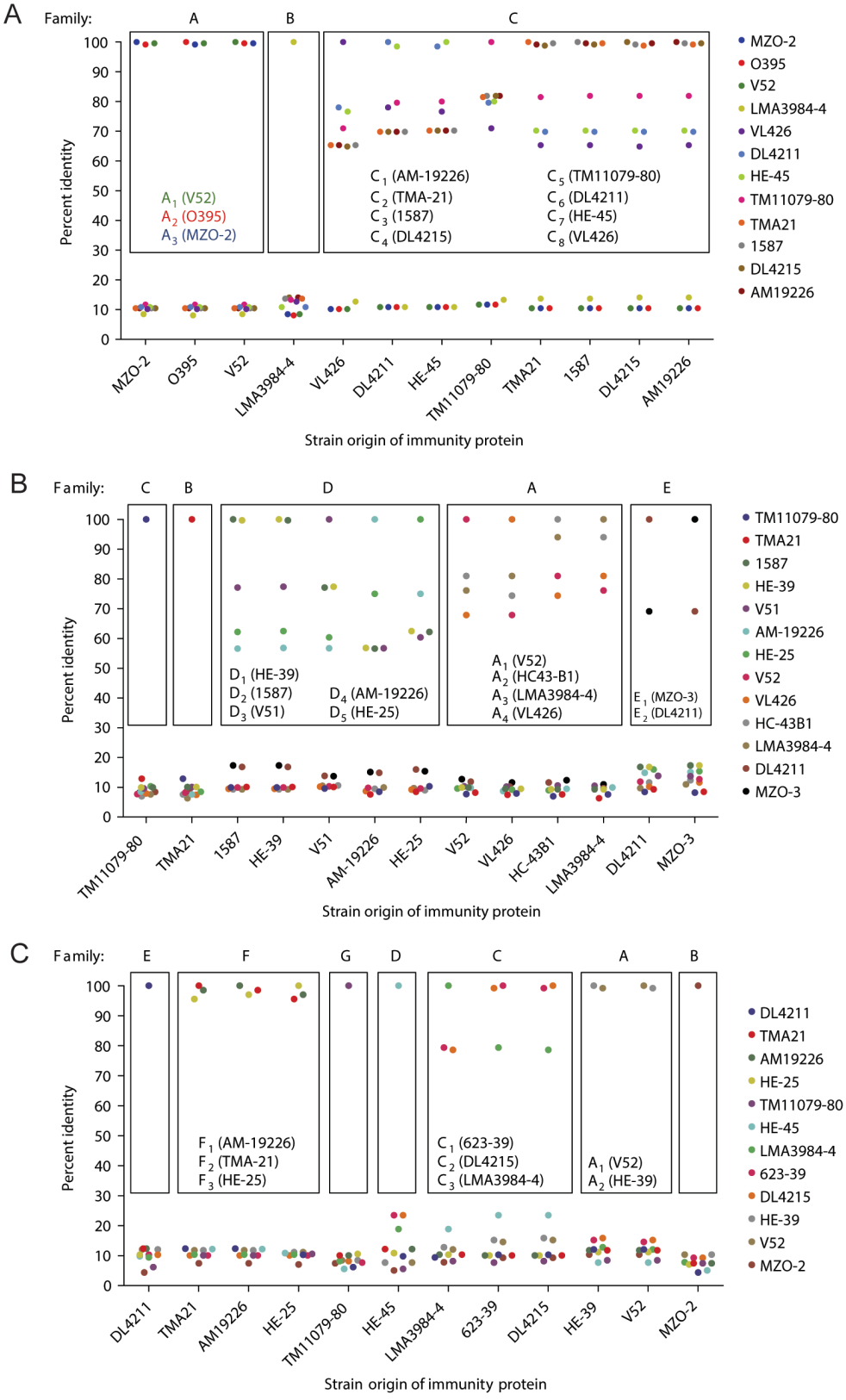

A
Family:
A
B
C
Percent identity
A1 (V52)
A2 (O395)
A3 (MZO-2)
C1 (AM-19226)
C2 (TMA-21)
C3 (1587)
C4 (DL4215)
C5 (TM11079-80)
C6 (DL4211)
C7 (HE-45)
C8 (VL426)
MZO-2
O395
V52
LMA3984-4
VL426
DL4211
HE-45
TM11079-80
TMA21
1587
DL4215
AM19226
Strain origin of immunity protein
B
Family:
C
B
D
A
E
D1 (HE-39)
D2 (1587)
D3 (V51)
D4 (AM-19226)
D5 (HE-25)
A1 (V52)
A2 (HC43-B1)
A3 (LMA3984-4)
A4 (VL426)
E1 (MZO-3)
E2 (DL4211)
TM11079-80
TMA21
1587
HE-39
V51
AM-19226
HE-25
V52
VL426
HC-43B1
LMA3984-4
DL4211
MZO-3
Strain origin of immunity protein
C
Family:
E
F
G
D
C
A
B
F1 (AM-19226)
F2 (TMA-21)
F3 (HE-25)
C1 (623-39)
C2 (DL4215)
C3 (LMA3984-4)
A1 (V52)
A2 (HE-39)
DL4211
TMA21
AM19226
HE-25
TM11079-80
HE-45
LMA3984-4
623-39
DL4215
HE-39
V52
MZO-2
Strain origin of immunity protein

Figure 4-3 Pairwise comparison of TsiV1, TsiV2 and TsiV3 immunity proteins encoded in the T6SS effector modules

The names of the strains that harbour the analysed immunity proteins in their auxiliary cluster 1 (a), auxiliary cluster 2 (b) or the large cluster (c) are shown on the x axis and in the legend on the right. Percent sequence identity between two pairwise-aligned amino-acid sequences is indicated on the y axis. Immunity proteins with more than 30% sequence identity were grouped into families (indicated by boxes). Subfamilies are indicated within the boxes: each dot indicates a single subfamily defined by an identical amino-acid sequence. One representative is shown for each subfamily. Reprinted by permission from McMillan Publishers Ltd: Nature Communications, Unterweger et al., 2014, copyright 2014.

To further differentiate between the sequences within one family that are all labeled with the same alphabetical letter like the ones from V52 and O395, an additional numeric character was devoted to the alphabetical letter like A_1 and A_2 to indicate that the sequences share more than 30% but less than 100% identity. Sequences that are 100% identical share the same letter and the same number. The same analysis was performed for sequences in auxiliary cluster 2 and the large cluster.

The labeling of these sequences with letters and numbers was then adapted as labeling for the whole effector module harboring the gene that encodes the respective immunity protein. For example, the immunity protein TsiV1 was devoted A1 in the analysis above. From hereon on we refer to the whole effector module encoding *tsiV1* as A1 _ _. Subsequently, multiple identically labeled effector modules contain genes that encode identical immunity proteins. Effectors of identically labeled effector modules can display varying degrees of polymorphisms (up to 1.6% difference among the amino acid sequences of effectors in the module in auxiliary cluster 1, up to 4.5% differences among effectors in the module in auxiliary cluster 2, up to 1.8% differences among effectors in the large cluster).

To determine the effector module set of an individual *V. cholerae* strain, we list the effector module families in the order of the gene clusters auxiliary cluster 1, auxiliary cluster 2 and the large cluster they are found in. For example, strain V52 harbors effector module A_1 in auxiliary cluster 1, effector module A_1 in auxiliary cluster 2 and effector module A_1 in the large cluster. Subsequently, the effector module set of V52 is referred to as $A_1A_1A_1$.

In summary, we made use of the similarities and differences between immunity proteins to classify effector modules and assign each *V. cholerae* strain a code representing its effector module set.

4.2.4 Strains with different effector module sets are incompatible

The effector module sets of *V. cholerae* strains revealed that some strains like V52 and C6706 harbor the same whereas other strains like V52 and AM-19226 harbor different effector module sets. We hypothesized that strains that harbor the same effector module set coexist and are compatible whereas strains that harbor different effector module sets outcompete each other and are thus incompatible.

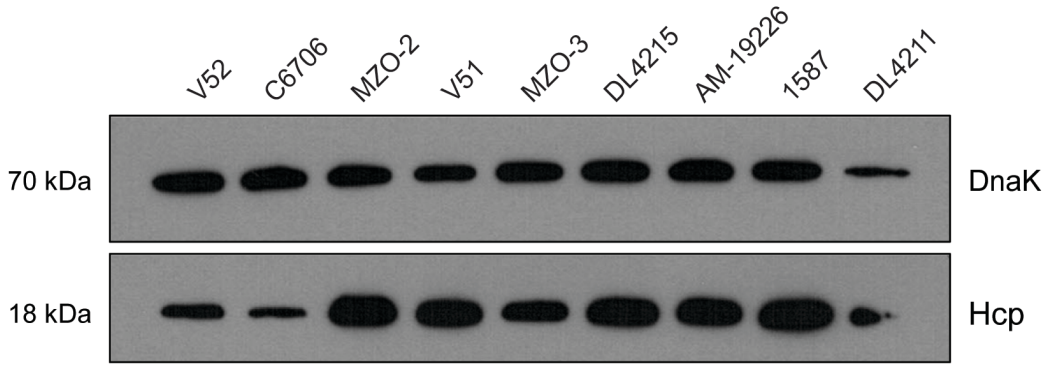


Figure 4-4 Hcp expression of indicated *V. cholerae* strains

Strains were spotted onto an LB agar plate at a density of 10^8 bacteria/25 µL spot and incubated for 4 h at 37ºC. Cells were harvested, lysed and subject to SDS-PAGE followed by western-blot analysis. Levels of Hcp and DnaK in whole cell lysates are shown. Reprinted by permission from McMillan Publishers Ltd: Nature Communications, Unterweger et al., 2014, copyright 2014.

To challenge our compatibility model, we co-incubated V52 ($A_1A_1A_1$ effector module set) with T6SS-active strains (Figure 4-4) harbouring the following effector module sets: $A_1A_1A_1$ (C6706), $A_3A_1B_1$ (MZO-2), $C_5D_3A_1$ (V51), $C_1E_1A_1$ (MZO-3), $C_4E_1C_2$ (DL4215), $C_3D_2C_1$ (1587), $C_1D_4F_1$ (AM-19226), or $C_6E_2E_1$ (DL4211) (Figure 4-5). In agreement with our hypothesis, all $A_1A_1A_1$ strains tested were able to coexist (Figure 4-5B), whereas non-$A_1A_1A_1$ *V. cholerae* were effectively outcompeted by wild-type V52 when

inoculated at a 1:1 ratio (Figure 4-5A). Control experiments carried out with a V52Δ*vasK* in-frame deletion mutant unable to translocate effectors confirmed that killing was T6SS-dependent (Figure 4-5A and B). To further test the compatibility model, the individual strains were also combined with each other in a killing assay (Figure 4-5 C). Strains that belong to the same compatibility groups turned out to co-exist whereas strains of differing compatibility groups turned out to kill each other. The O1 serogroup strain NCTC8457 did not behave as expected. Despite belonging to the AAA compatibility group, NCTC8457 does kill the strain V52 of the AAA compatibility group.

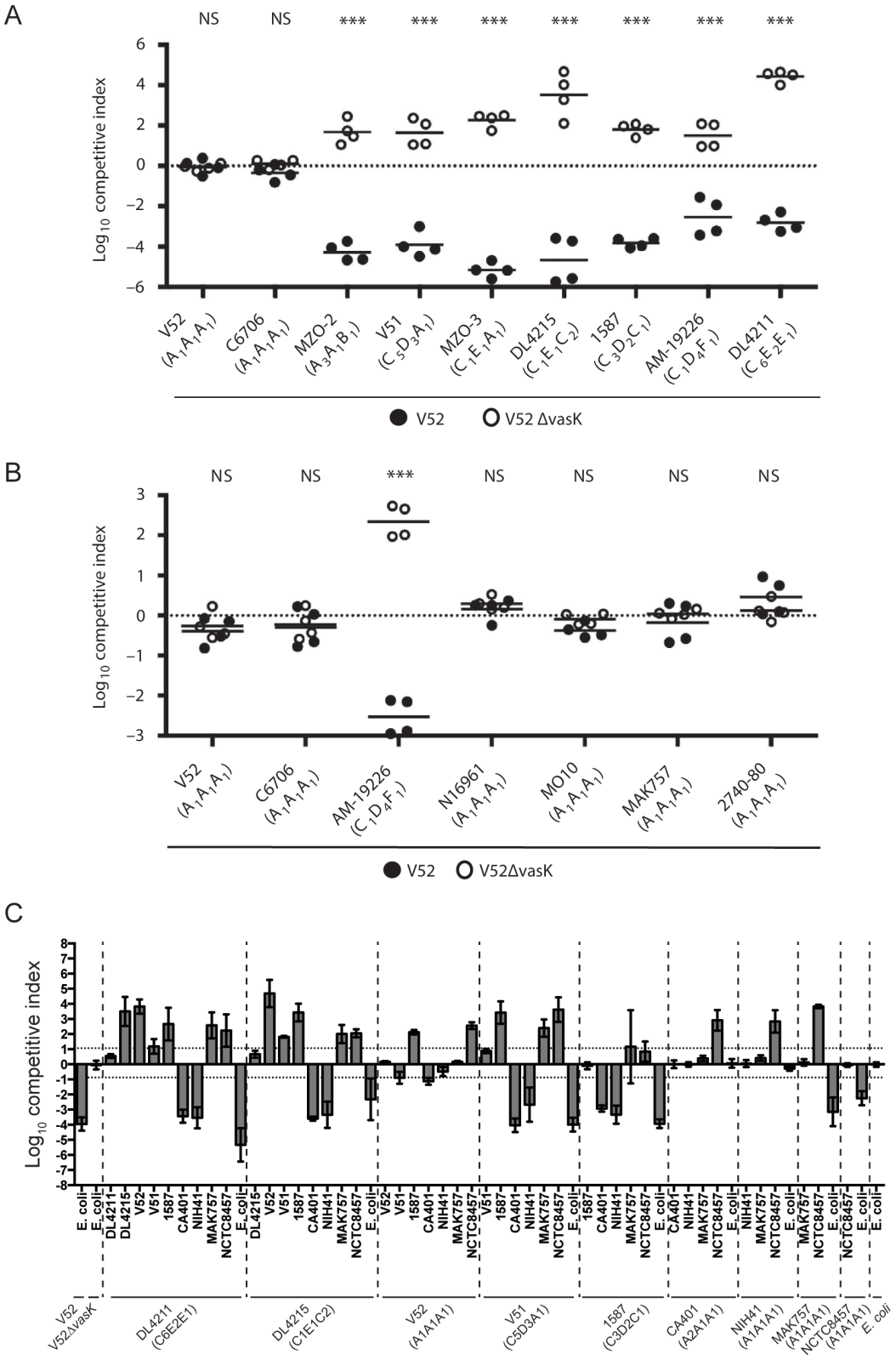

A
NS
NS
***
***
***
***
***
***
***
Log10 competitive index
6
4
2
0
−2
−4
−6
V52 (A1A1A1)
C6706 (A1A1A1)
MZO-2 (A3A1B1)
V51 (C5D3A1)
MZO-3 (C1E1A1)
DL4215 (C1E1C2)
1587 (C3D2C1)
AM-19226 (C1D4F1)
DL4211 (C6E2E1)
V52
V52 ΔvasK
B
NS
NS
***
NS
NS
NS
NS
Log10 competitive index
3
2
1
0
−1
−2
−3
V52 (A1A1A1)
C6706 (A1A1A1)
AM-19226 (C1D4F1)
N16961 (A1A1A1)
MO10 (A1A1A1)
MAK757 (A1A1A1)
2740-80 (A1A1A1)
V52
V52ΔvasK
C
Log10 competitive index
8
7
6
5
4
3
2
1
0
-1
-2
-3
-4
-5
-6
-7
-8
E. coli
DL4211
DL4215
V52
V51
1587
CA401
NIH41
MAK757
NCTC8457
V52
V52ΔvasK
DL4211 (C6E2E1)
DL4215 (C1E1C2)
V52 (A1A1A1)
V51 (C5D3A1)
1587 (C3D2C1)
CA401 (A2A1A1)
NIH41 (A1A1A1)
MAK757 (A1A1A1)
NCTC8457 (A1A1A1)
E. coli

Figure 4-5 T6SS effector modules govern compatibility

T6SS diversity determines competition (A) and coexistence (B). Wild-type or *vasK*-deficient V52 was mixed with indicated *V. cholerae* strains; the mixture was incubated on nutrient agar for 4 h, and survival was enumerated by plating survivors on appropriate selective plates. The results of two independent experiments performed in duplicate are shown. Horizontal bars indicate the arithmetic mean of log-transformed data. Stars indicate statistical significance (unpaired, two tailed Student's t-test: ***P<0.0005; NS, P>0.05). Reprinted by permission from McMillan Publishers Ltd: Nature Communications, Unterweger et al., 2014, copyright 2014. (C) Killing assays between the indicated strains as described above. The arithmetic mean ± SD of log-transformed data from two independent experiments performed in duplicate are shown.

These results suggest that *V. cholerae* strains with the same effector module set co-exist and are thus compatible whereas strains that harbor different effector module sets outcompete each other in a T6SS-dependent manner, and are thus incompatible. Subsequently, we propose the analysis of effector module sets as a method to identify compatible *V. cholerae* strains, also referred to as strains of one compatibility group. We observed an exception to the compatibility rule. The O1 serogroup strain NCTC8457 belongs to the AAA compatibility group but does kill another AAA strain. Further experiments need to reveal if this is due to a T6SS-dependent or T6SS-independent killing mechanism. The separation of strains into compatible and incompatible strains is only relevant when bacteria have an active T6SS and under conditions that permit T6SS-dependent competition, like an agar plate and not liquid culture.

4.2.5 Bi-directional killing between T6SS-active strains of different compatibility groups

To further characterize the ability of two strains with different effector module sets to fight against each other in a T6SS-dependent manner, we tested the ability of the strains V52 and AM-19226 to kill each other.

Both strains have an active T6SS under lab conditions as shown by the secretion of Hcp into culture supernatants (Figure 4-6A). Analysis of the anti-prokaryotic activity of the two strains in killing *E. coli* shows that V52 and AM-19226 engage in T6SS-mediated killing.

When mixing bacteria of the strains V52 and AM-19226 in a killing assay at equal ratios following a 4-h incubation period, the number of surviving V52 is 10000 fold higher than the number of surviving AM-19226 (Figure 4-6C). When mixing *vasK*-deficient mutants of both strains, surviving bacteria of the strains V52 and AM-19226 are rescued at equal numbers indicating that any differences in the ratio between the two strains can be attributed to T6SS-mediated killing only. When testing V52 or AM-19226 against a *vasK*-deficient mutant of the respective other strain, AM-19226 is killing V52Δ*vasK* by 2 logs whereas V52 is killing AM-19226Δ*vasK* by 5 logs. These results show a difference in the ability of the individual strains V52 and M-19226 to kill *V. cholerae* that was not observed for the killing of *E. coli* (Figure 4-6B).

In summary, we observe differences between *V. cholerae* strains to kill *E. coli* or each other. In addition, our results indicate that intraspecific competition between two T6SS active strains is bidirectional, because, even though bacteria of the strain V52 are surviving at 10,000 fold higher numbers than AM-19226, bacteria of the strain V52 survive by one more log after exposure to AM-19226Δ*vasK* than to AM-19226 wild-type.

To test the influence of the input ratio between V52 and AM-19226 on the outcome of the killing assay, we performed killing assays on V52 and AM-19226 mixed at equal total numbers but different input ratios ranging from 1/100 to 100/1 (Figure 4-6D).

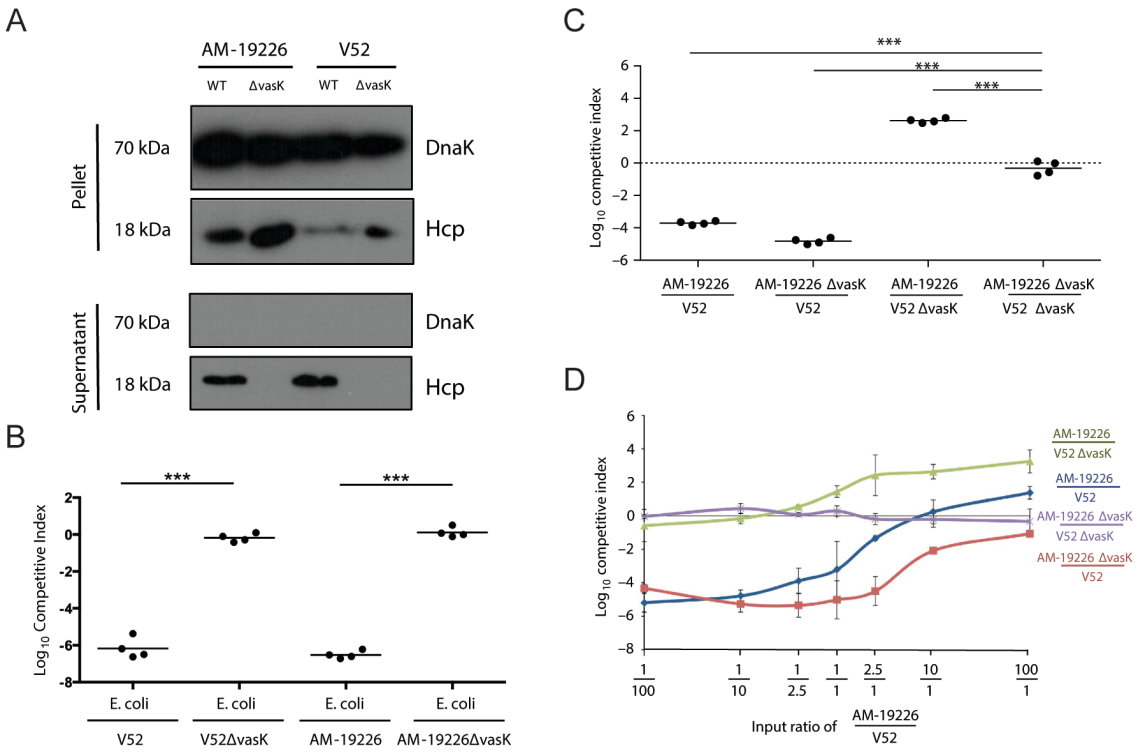


Figure 4-6 T6SS-mediated killing between two strain of different compatibility groups

(A) Strains were grown in liquid LB broth to mid-logarithmic phase under laboratory conditions. After centrifugation, pellets and supernatants were analysed by western-blotting with polyclonal rabbit anti-Hcp antiserum and monoclonal mouse anti-DnaK antibody. (B) A competition assay of *V. cholerae* strains V52 and AM-19226 (wild-type and Δ*vasK* mutants) against *E. coli* at a 10:1 ratio. The results of two independent experiments that were each performed in duplicate are shown. Horizontal bars represent the arithmetic mean of log-transformed data. Error bars indicate the standard deviation. Stars indicate statistical significance (unpaired, two-tailed student's t-test: t6=19.75, P<0.0001 and t6=38.55, P<0.0001). (C-D) Wild-type or *vasK*-deficient V52 was mixed with wild-type or *vasK*-deficient AM-19226 at a ratio of 1:1 (C) or at various ratios as shown on the x axis (D). The logarithm of the competitive index of AM-19226 over V52 after 4 h coincubation is shown on the y axis. Horizontal bars indicate the arithmetic mean of log-transformed data based on two independent experiments, each performed in duplicate. Stars indicate statistical significance (unpaired, two-tailed Student's t-test: ***P<0.0005). Error bars indicate the s.d. of log-transformed data.

Independent of the input ratio, AM-19226Δ*vasK* mixed with V52Δ*vasK* was always recovered at around a ratio 1:1 after 4 hour co-incubation. When mixing bacteria of both wild-type strains, more V52 was recovered than AM-19226 until an input ratio of about 1/5 when they were recovered at equal ratios. For ratios smaller than 1/5, more

about 1/5 when they were recovered at equal ratios. For ratios smaller than 1/5, more AM-19226 was recovered. When analyzing unidirectional killing, the ratio of AM-19226 over V52Δ*vasK* needed to be at least 1/2.5 for AM-19226 to be recovered at higher numbers. Bacteria of the strain V52 were recovered at higher numbers than AM-19226Δ*vasK* when mixed with each other at a ratio as low as 1/10.

These results suggest that the input ratio influences the competitive outcome of T6SS-mediated intraspecific competition. During the analysis of unidirectional killing, we observed strain-specific differences in the magnitude of T6SS-mediated killing and the ratios at which a given strain was able to outcompete another strain.

4.2.6 T6SS effector modules show hallmarks of horizontal gene transfer

In the previous experiments, we established a concept of compatibility and incompatibility between *V. cholerae* strains based on T6SS-mediated killing. We then hypothesized that the ability or inability to co-exist with each other is reflected in the phylogeny of compatible and incompatible strains.

To test this hypothesis we built a phylogenetic tree of the 37 analyzed *V. cholerae* strains based on six polymorphic housekeeping genes (Boucher et al., 2011) (Figure 4-7). We added the information on the compatibility group for each strain and compared compatibility group membership to relatedness between strains. Strains like Amazonia and TM11079-80 derive from the same recent common ancestor and are closely related. They belong to the same compatibility group. The O1 serogroup strain 12129(1), which was isolated form an environmental sample in Australia in 1985, and the O1 serogroup strain LMA3984-4, which was isolated from the environment in Brazil in 2007, are also derived from the same recent common ancestor but represent different compatibility groups. This observation led us to conclude that two *V. cholerae* strains can harbor different effector module sets of which maximally one can be derived from the recent common ancestor of both strains.

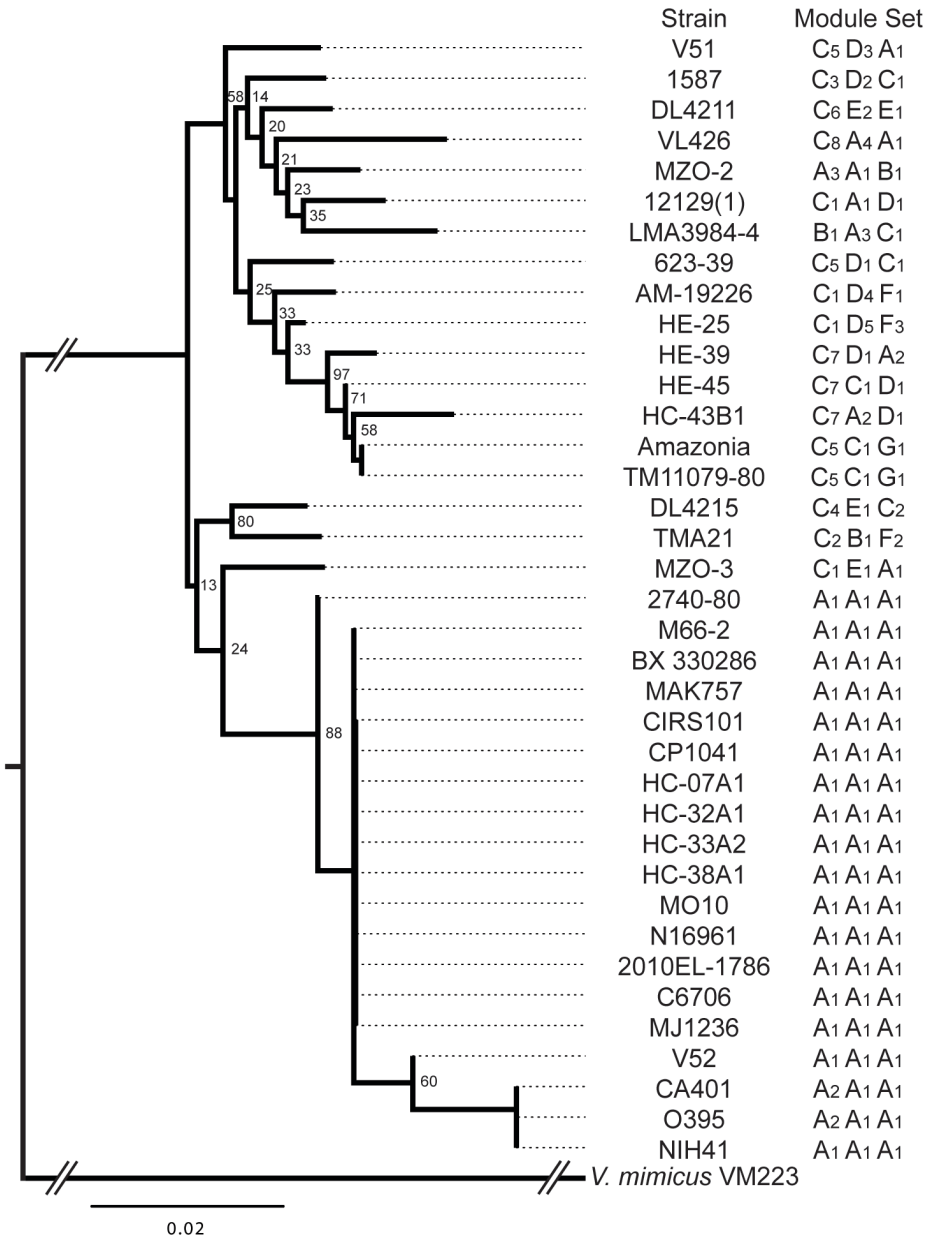


Figure 4-7 Phylogenetic relationship of *V. cholerae* strains with a variety of T6SS effector module sets.

A neighbour-joining phylogenetic tree (bootstrap=100) of the 37 *V. cholerae* strains analysed based on the six housekeeping genes *adk*, *gyrB*, *mdh*, *recA*, *pgi* and *rpoB* is shown. Effector module sets indicating the family of the effector module in auxiliary cluster 1, auxiliary cluster 2 and the large cluster are indicated next to the strain name. Reprinted by permission from McMillan Publishers Ltd: Nature Communications, Unterweger et al., 2014, copyright 2014.

Previous analysis of the GC content of effector modules suggests that these modules are subject to horizontal gene transfer. We hypothesized that effector modules are also exchanged between *V. cholerae* strains. Therefore, we analyzed similar effector modules in distantly related strains, like the effector module A3 in the large cluster of the strains V51 and VL426. Alignment of the immunity-protein encoding genes shows that the two sequences are 100% identical whereas the concatenated sequences of the house-keeping genes differ by 1.5% (Figure 4-8 A and B). This finding suggests that the effector module A3 rather got acquired by horizontal gene transfer than inherited from a common ancestor. Additional examples are found for effector modules in auxiliary cluster 1 and auxiliary cluster 2, as shown by the high similarity of *tsiV1* sequences independent of the phylogenetic relation between the strains whose genome they are contained in Figure 4-8C.

Altogether, these results suggest that effector modules are acquired by horizontal gene transfer resulting in incompatible strains derived from the same recent common ancestor.

Figure 4-8 Alignment of immunity protein encoding genes and house keeping genes.

(A) Alignment of the concatenated nucleotide sequences of the genes *adk, gyrB, mdh,* recA*, pgi* and *rpoB*. Green bars correspond to 100%, yellow bars to less than 100% identity between the nucleotides of the two compared sequences. The strain names are indicated on the left. (B) Alignment of the nucleotide sequences of *tsiV3A1* of V51 and VL426. (C) Phylogenetic trees of *tsiV1*A, *tsiV2*A1, and *tsiV3*A of indicated strains. Reprinted by permission from McMillan Publishers Ltd: Nature Communications, Unterweger et al., 2014, copyright 2014.

4.2.7 Pandemic strains contain T6SS effector module sets of the same family and are compatible

19 of the 37 analyzed *V. cholerae* strains belong to the compatibility group AAA. Among them are pandemic strains like N16961 and C6706. We hypothesized that compatibility group membership separates pandemic from non-pandemic strains.

Therefore, we analyzed the *V. cholerae* strains for their repertoire of virulence factors. The presence or absence of Vibrio pathogenicity island 1 (VPI-1) which contains

the Tcp-encoding gene and the presence or absence of CTX-Φ that harbors the CTX-encoding genes were analyzed. Analysis of the strains and the virulence factors they encode showed that 18 strains encode VPI-1 and CTX-Φ, 2 strains encode only VPI-1 and 17 strains encode neither (Figure 4-9). All of the strains that encode virulence factors and are pandemic belong to the AAA compatibility group. Two AAA strains do not encode both virulence factors. The strain V51 of the compatibility group CDA encodes virulence factors but is not a pandemic isolate.

We therefore concluded that all pandemic strains belong to the AAA compatibility group and are thus expected to co-exist. We further noticed that pandemic strains of the classical biotype belong to the $A_2A_1A_1$ compatibility group whereas the El Tor pandemic strains belong to the $A_1A_1A_1$ compatibility group. An in-depth analysis following up on this observation is provided in chapter 6.

Strain	Module Set	VPI-1/CTX-Φ
V51	C_5 D_3 A_1	+/+
1587	C_3 D_2 C_1	-/-
DL4211	C_6 E_2 E_1	-/-
VL426	C_8 A_4 A_1	-/-
MZO-2	A_3 A_1 B_1	-/-
12129(1)	C_1 A_1 D_1	-/-
LMA3984-4	B_1 A_3 C_1	-/-
623-39	C_5 D_1 C_1	-/-
AM-19226	C_1 D_4 F_1	-/-
HE-25	C_1 D_5 F_3	-/-
HE-39	C_7 D_1 A_2	-/-
HE-45	C_7 C_1 D_1	-/-
HC-43B1	C_7 A_2 D_1	-/-
Amazonia	C_5 C_1 G_1	-/-
TM11079-80	C_5 C_1 G_1	-/-
DL4215	C_4 E_1 C_2	-/-
TMA21	C_2 B_1 F_2	-/-
MZO-3	C_1 E_1 A_1	-/-
2740-80	A_1 A_1 A_1	+/-
M66-2	A_1 A_1 A_1	+/-
BX 330286	A_1 A_1 A_1	+/+
MAK757	A_1 A_1 A_1	+/+
CIRS101	A_1 A_1 A_1	+/+
CP1041	A_1 A_1 A_1	+/+
HC-07A1	A_1 A_1 A_1	+/+
HC-32A1	A_1 A_1 A_1	+/+
HC-33A2	A_1 A_1 A_1	+/+
HC-38A1	A_1 A_1 A_1	+/+
MO10	A_1 A_1 A_1	+/+
N16961	A_1 A_1 A_1	+/+
2010EL-1786	A_1 A_1 A_1	+/+
C6706	A_1 A_1 A_1	+/+
MJ1236	A_1 A_1 A_1	+/+
V52	A_1 A_1 A_1	+/+
CA401	A_2 A_1 A_1	+/+
O395	A_2 A_1 A_1	+/+
NIH41	A_1 A_1 A_1	+/+

Figure 4-9 Overview over *V. cholerae* strains, their T6SS effector modules and VPI-1/CTX-Φ

Effector module sets indicating the family of the effector module in auxiliary cluster 1, auxiliary cluster 2 and the large cluster, and VPI-1/CTX-Φ acquisition are indicated next to the strain name. Strains that harbour the AAA effector module set are highlighted in the blue box.

4.3 Discussion

Analysis of 37 *V. cholerae* strains and their T6SS gene clusters revealed basic principles of intraspecific competition among *V. cholerae* strains (Figure 4-10). We observed a variety of putative T6SS effector and immunity proteins. These proteins are encoded in genes of effector modules that are likely to be acquired by horizontal gene transfer. We can use the effector modules to assign strains to compatibility groups of strains which can co-exist and between which strains outcompete each other, as tested experimentally. All analyzed pandemic *V. cholerae* strains were shown to belong to the same compatibility group and are thus expected to be compatible among each other but not with non-pandemic strains.

The T6SS effectors with anti-prokaryotic activity TseL, VasX and VgrG-3 of the *V. cholerae* strain V52 have previously been characterized in our lab and by others (Brooks et al., 2013; Dong et al., 2013; Miyata et al., 2013; Russell et al., 2013). Beyond these three effectors, this study revealed that *V. cholerae* strains harbour a variety of putative T6SS effectors encoded in the same position and with a predicted diversity of effector activities. Biochemical analysis similar to the ones described for the pore-forming effector VasX (Miyata et al., 2013), the peptidoglycan-degrading effector VgrG-3 (Brooks et al., 2013; Russell et al., 2011) and the lipase effector TseL (Russell et al., 2013) are required to experimentally test the anti-prokaryotic activity and substrate specificities of these putative effectors. The bioinformatics analysis predicted one effector activity like peptidoglycan degradation for multiple putative T6SS effectors. Because their amino acid sequences differed from each other over 70 percent, we propose that individual effectors have different substrate specificities as previously described for peptidoglycan- and lipid-degrading effectors in other bacterial species (Russell et al., 2011; Russell et al., 2013). Once the individual effectors are characterized, I suggest comparing their substrate specificities and target range to learn more about the competitive advantage or disadvantage a particular effector is providing to the bacterium it is found in.

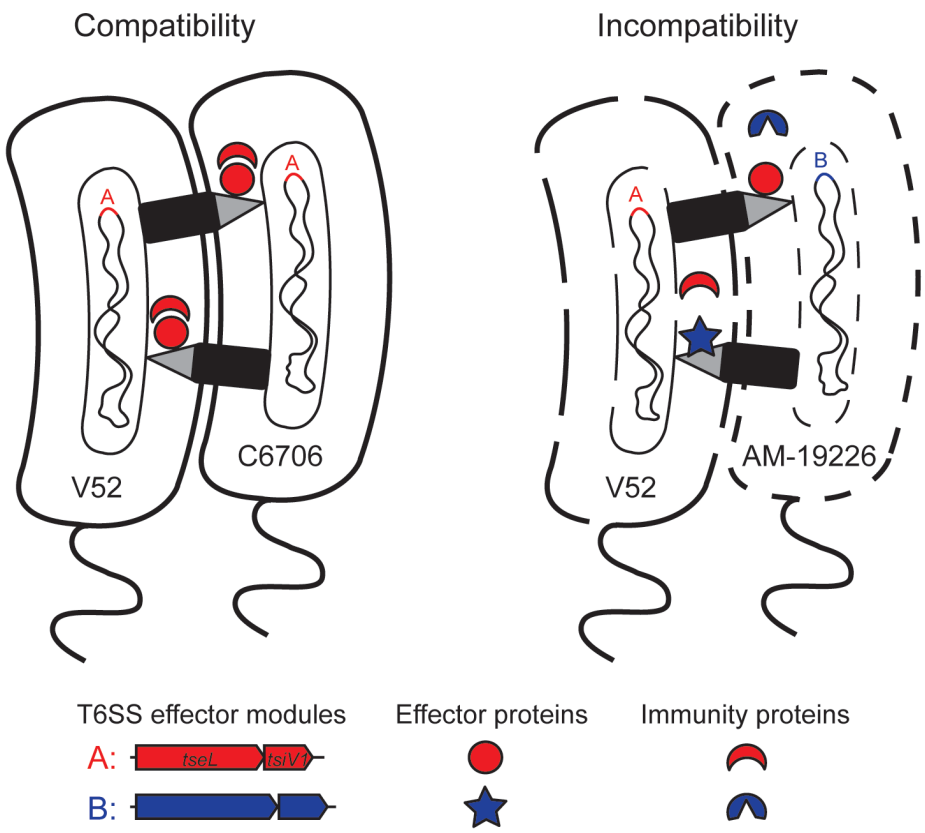


Figure 4-10 Effector modules determine compatibility and incompatibility between *V. cholerae* strains

Graphical depiction of two *V. cholerae* bacteria from two different indicated strains. The outer membrane, inner membrane and chromosomal DNA are depicted. The T6SS is illustrated with a black sheath and a grey tip.

In addition to the variety of putative effector proteins, this study also revealed a variety of putative T6SS immunity proteins. We observed that effector modules that encode genes for the same effector protein, encode genes for the same immunity protein whereas effector modules that encode genes for different effector proteins also encode genes for different immunity proteins. This co-segregation between effector and immunity protein-encoding genes suggests that the described proteins are specific to their cognate effector similar to the cognate inhibition of TsiV1, TisV2 and TsiV3 to TseL, VasX and VgrG-3, respectively (Unterweger et al., 2014). The function of the variety of putative immunity proteins found among the 37 strains to specifically inhibit their cognate effector has to be further tested experimentally. Our comparison between the

amino acid sequences of multiple immunity proteins showed that the sequences differ on multiple levels from each other. Some sequences differ by more than 70 percent amino acid sequence identity, others by 30, 20 or 10 percent. To further investigate the interaction between these immunity proteins and effectors, we started a collaboration with Alexei Savchenko (University of Toronto) to determine the structural basis for compatibility between strains. We are currently co-expressing proteins to overcome the hurdle of unstable and insoluble proteins when expressed at high quantities. Solving the crystal structures of immunity proteins in complex with effector proteins will reveal the interaction sites between the two proteins. Subsequently, we will be able to determine which differences in the amino acid sequences between proteins affect the interaction between the effector and immunity protein. Determining the crystal structures of the immunity proteins in complex with their cognate effector will also provide further insight into the mode of inhibition. TsiV3 inhibits VgrG-3 by competing with the substrate for binding in the active site. Other than such competitive inhibition, it could be imagined that immunity proteins also inhibit effectors by allosteric inhibition. Observations by Zhang and colleagues indicate that dimerization between TsiV3 immunity proteins was required for protection from VgrG-3 (Zhang et al., 2014). Mutants of TsiV3 with amino acid substitutions at positions required for the dimerization failed to protect from VgrG-3 (Zhang et al., 2014).The diversity of T6SS effector and immunity proteins is associated with the ability of numerous *V. cholerae* strains to engage in T6SS-mediated intraspecific competition. The ability to engage in intraspecific bacterial killing, which we originally observed using environmental isolates from the Rio Grande (Unterweger et al., 2012), is observed among strains independent of their serogroup, time or place of isolation. Analysis of compatible and incompatible strains in respect to their phylogenetic relationship to each other and the virulence factors they encode indicate that the T6SS is an important feature of bacteria within the species that might have contributed to the phylogeny and virulence of the species as we know it today. In the phylogenetic tree, we observe examples for closely related strains deriving from the same common ancestor to belong to the same compatibility group and distantly related strains of the same recent common ancestor to belong to a different compatibility group. Analysis of the strains and their virulence factors indicates that pandemic strains all co-exist among each other but

not with non-pandemic strains. T6SS-mediated killing might be a mechanism to select for pandemic strains and could have important implications during infection. Humans likely take up a mixed inoculum of pandemic and non-pandemic strains (Faruque et al., 2004). To get maximal access to limited resources in the intestine, pandemic strains might kill non-pandemic strains and successfully establish an infection. This model is supported by a study of Stine and colleagues, in which toxigenic and non-toxigenic strains were isolated from the environment, but only toxigenic strains from cholera patients (Stine et al., 2008). I propose an experiment to test T6SS-mediated intraspecific competition *in vivo*. Within the small intestine of infant rabbits, T6SS-mediated competition between C6706 wild-type and C6706Δ*tsiV3* has been reported (Fu et al., 2013). Providing the infant rabbit with combinations of C6706 wild-type - AM-19226Δ*vasK* and C6706Δ*vasK* - AM-19226Δ*vasK* will reveal if bacteria of the strain C6706 colonize the infant rabbit better when equipped with a functional T6SS to kill other *V. cholerae* bacteria of the inoculum.

Considering the importance of T6SS-mediated intraspecific competition for the species *V. cholerae*, we propose a labeling system of compatibility groups based on the effector module set an individual strain harbors in its T6SS gene clusters. We tested a selection of strains for their ability to co-exist or to compete with the strain V52 of which the T6SS effectors are characterized. All strains turned out to co-exist with bacteria of the strain V52 that were predicted to be compatible because they harbor effector modules sets of the same family. Strains that were predicted to be incompatible with the strain V52 competed with bacteria of the strain V52 in a killing assay. These experiments show a correlation between strains that encode the same effector modules to be compatible and strains that encode different effector module sets to outcompete each other. More strains have to be tested to further validate this model and additional strains deficient in their effector modules need to be created to test a general causative relationship between effector modules and compatibility. The proposed model of compatibility should be further extended to T6SS-effectors encoded in genes outside of the here analyzed effector modules and to T6SS-independent anti-prokaryotic factors like observed for bacteria of the strain DL4215 (Unterweger et al., 2012) to provide a robust prediction for compatibility and incompatibility between bacteria. A uniform labeling system for

compatibility groups as proposed in here could become a powerful tool. We observe that pandemic strains are associated with the AAA compatibility group. The advantage for pandemic strains of having the particular set of the three effectors TseL, VasX and VgrG-3 remains to be elucidated. Possible advantages could be a substrate specificity that allows pandemic strains to kill a wide range of other bacteria, including non-pandemic *V. cholerae* strains and bacteria of the intestinal microbiota. Membership to this compatibility group could be used as an indicator for a pandemic strain during outbreak management or clinical diagnostics. The association between pandemic strains and a specific set of T6SS effectors could also be used to develop a therapeutic strategy targeting the T6SS effectors of pandemic strains.

Compatibility groups were determined on the basis of effector modules. Our results suggest that these effector modules are mobile genetic elements that can be acquired by horizontal gene transfer. Strong indicators are the difference in GC content between effector modules and the surrounding genes, and the distribution of effector modules among strains independent of their phylogenetic relationship. The exchange of effector modules between *V. cholerae* strains remains to be shown experimentally. I propose an experiment in which genomic DNA that contains an effector module marked with an antibiotic resistance cassette is provided to naturally competent *V. cholerae* bacteria that harbor a different effector module in this locus. A copy of the immunity protein-encoding gene will be inserted at a separate site on the chromosome to prevent that the bacterium that is taking up a new effector module is getting killed by parental kin. Selective plates supplemented with the antibiotic could be used to isolate bacteria that acquired a new effector module. The respective locus in the genome could be sequenced to test if the original effector modules actually got replaced by the effector module with the antibiotic resistance cassette. In additional experiments, the molecular mechanism of effector module exchange and the frequency of this event could be further examined. Ultimately, I suggest comparing the competitive fitness of two strains that are genetically identical except for their effector modules to analyze the benefit of an effector module exchange for the bacterium during the encounter of a competitor.

The alignment of the T6SS gene clusters among the analyzed 37 *V. cholerae* strains revealed a high conservation of genes encoding structural and regulatory

components. In contrast, the effector modules encode very diverse effector proteins that lack an obvious signal sequence. The ability of the structurally conserved T6SS to translocate up to 105 theoretical combinations of three T6SS effectors provides the basis for T6SS-mediated intraspecific competition. The underlying mechanism still remained unknown at this point.

Chapter 5

Type VI secretion system adaptor protein 1 (Tap-1) provides mechanistic basis for secretion of diverse cargo effectors

Portions of this chapter are accepted for publication:

Unterweger D., Kostiuk, B., Oetjengerdes, R., Wilton, A., Diaz-Satizabal, L., Pukatzki, S., Chimeric adaptor protein Tap-1 translocates diverse type VI secretion system effectors in *Vibrio cholerae*. EMBO J accepted.

Ben Kostiuk performed the killing assay shown in Figure 5-3B. Rina Oetjengerdes performed the killing assay shown in Figure 5-8. Ashley Wilton performed the killing assay shown in Figure 5-2 E. Stefan Pukatzki performed the expression analysis of the VgrG-1 constructs. Nikki Atanasova performed the killing assay shown in Figure 5-3C.

5 Type VI secretion system adaptor protein 1 (Tap-1) provides mechanistic basis for secretion of diverse cargo effectors

5.1 Introduction

V. cholerae strains harbor a variety of T6SS effectors that enable these strains to engage in intraspecific competition (Unterweger et al., 2014). Effector proteins are encoded in mobile genetic elements called effector modules that are likely exchanged between strains via horizontal gene transfer. Enzymatic activity and the mode of translocation with the T6SS apparatus differ between individual effector proteins. Effector domains can be part of VgrG-proteins like the C-terminal extension of VgrG-3 with peptidoglycan-degrading activity (Brooks et al., 2013; Dong et al., 2013). Cargo effectors like TseL and VasX are neither part of VgrG proteins nor do they contain a PAAR domain to directly interact with VgrG proteins. Despite the lack of an obvious signal sequence among cargo effectors, the secretion system apparatus translocating these effectors is conserved (Unterweger et al., 2014). The molecular mechanism by which cargo effectors are translocated with the T6SS was unknown at the time we published our results on the diversity of *V. cholerae* T6SS effectors (Unterweger et al., 2014).

Here, I am identifying the type VI secretion system adaptor protein 1 (Tap-1) and describe its requirement for the interaction between the cargo effector TseL and VgrG-1. Subsequently, TseL requires Tap-1 for secretion. Strains with differing cargo effectors harbour chimeric adaptor proteins that facilitate the translocation of their cognate effector. I further describe a putative recombination site within *tap-1* that might enable the exchange of effector modules in auxiliary cluster 1.

5.2 Results

5.2.1 T6SS gene clusters encode proteins of the DUF4123 superfamily

The anti-prokaryotic T6SS effector TseL is encoded in the auxiliary cluster 1 of pandemic *V. cholerae* O1 serogroup strains and other clinical isolates such as the O37 serogroup strain V52 (Dong et al., 2013). We set out to elucidate the mechanism by which the T6SS effector TseL is translocated into neighbouring cells. We took a candidate approach to decipher the molecular events that lead to TseL translocation. First, we evaluated the contribution of the uncharacterized ~ 33 kDa-protein Tap-1, encoded immediately upstream of *tseL* (Figure 5-1A). BLAST (Altschul et al., 1990) analysis revealed that Tap-1 belongs to a superfamily of proteins with the domain of unknown function, DUF4123 (Figure 5-1B). This domain is about 120 amino acids in length and contains multiple conserved motifs: YLLLD, SPYxxLVxL, HLRxLLxV and LFRFYDPxVL (Figure 5-1C). DUF4123-containing proteins are found in over 100 bacterial species. Five representative species with DUF4123-encoding genes as part of their T6SS clusters are depicted in Figure 5-1E. DUF4123-containing proteins often contain additional conserved domains, including Fha, ZapA and PWWP domains, all of which are involved in protein-protein interactions (Figure 5-1E). Genes encoding the DUF4123 superfamily proteins are commonly found downstream of *vgrG* or putative effector-encoding genes in predicted T6SS gene clusters (Figure 5-1D).

Some species encode multiple proteins of the DUF4123 superfamily in their genomes. For example, in addition to Tap-1, *V. cholerae* strain N16961 encodes VasW (VCA0019) (Figure 5-1D). We previously demonstrated that VasW is necessary for secretion of and killing mediated by VasX, the pore-forming effector encoded directly downstream of *vasW* (Miyata et al., 2013). Even though Tap-1 and VasW belong to the same superfamily, the two proteins share less than 20% identity in their amino acid sequences (Figure 5-1F).

In summary, *tap-1* encodes a protein of the DUF4123 superfamily that is found in T6SS gene clusters of many bacterial species. Its conserved nature and physical proximity to T6SS effectors suggests a conserved role for Tap-1 in T6SS function.

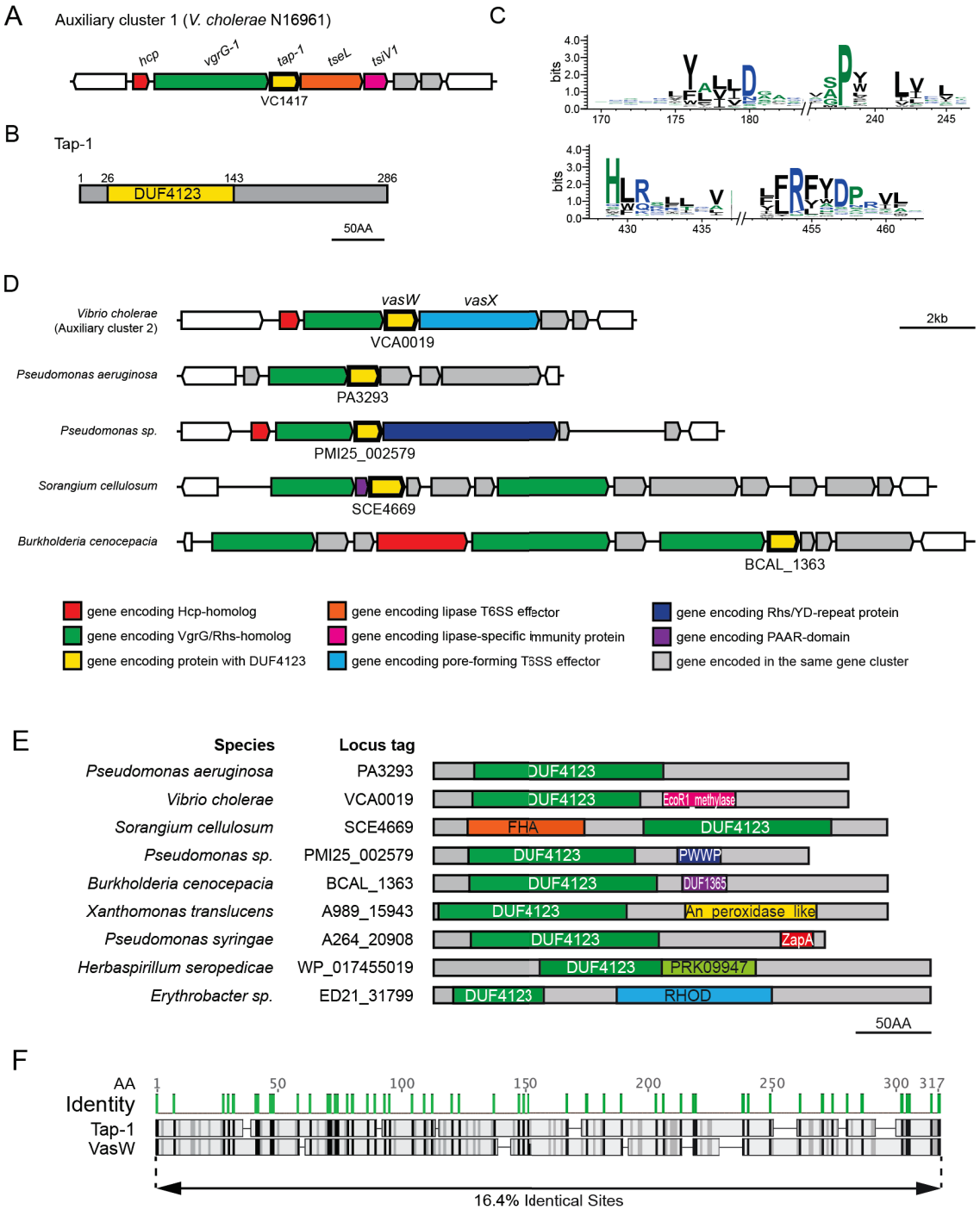


Figure 5-1 Tap-1 belongs to the superfamily of DUF4123 proteins.

Figure 5-1 Tap-1 belongs to the superfamily of DUF4123 proteins

(A) Graphical depiction of the T6SS auxiliary gene cluster 1 of *V. cholerae* strain N16961. Gene names are indicated above each gene. *Tap-1* (yellow) is outlined in bold and its locus tag is indicated below the gene. A legend for colour coding is shown below (C). (B) Graphical depiction of Tap-1 indicating the domain with homology to DUF4123. (C) WebLogo of the 4 characteristic motifs within DUF4123, from a sequence alignment among the 100 most diverse members of the DUF4123 superfamily. Colour of the letter indicates hydrophilic (blue), neutral (green) or hydrophobic (black) character of the amino acid. Value on the y-axis indicates sequence conservation at a particular site on the x-axis. If multiple amino acids are shown at one position, proportional height indicates relative frequency. (D) Graphical depiction of DUF4123-containing gene clusters drawn to scale. The bacterial species encoding the genes is indicated on the left. All genes encoding proteins of the DUF4123 superfamily (yellow) are outlined in bold and labeled with their locus tag. Legend indicates function predicted by BLAST. Genes of unknown function are grey. Genes outside the operon are white. (E) Cartoon (drawn to scale) of nine members of the superfamily from various bacterial species, showing the modular domain architecture of DUF4123 superfamily proteins that combine various domains with DUF4123. Conserved domains are coloured. (F) Alignment of the amino acid sequences of Tap-1 and VasW.

5.2.2 Tap-1 is required for TseL translocation

Tap-1 is encoded upstream of *tseL*. Based on their close proximity, we investigated whether Tap-1 is required for TseL secretion and subsequent TseL-mediated killing. To determine the secretion requirements for TseL, V52 and an isogenic *tap-1* mutant were maintained in LB broth until they reached the mid-logarithmic phase of growth. TseL was found in the pellet and supernatant from V52, but was absent from the supernatant of the V52 mutant lacking *tap-1* (Figure 5-2A). *In trans* complementation of the *tap-1* deletion restored detection of TseL in the supernatant. As expected, DnaK – a cytoplasmic protein used as a lysis control – was found in the pellet but not supernatant. We conclude that Tap-1 is required for T6SS-mediated translocation of TseL. To determine if the secretion defect of the *tap-1* mutant is specific for TseL or affects T6SS-mediated secretion universally, we analyzed Hcp secretion in V52Δ*tap-1,* as Hcp secretion is the hallmark of a functional secretion system. We observed that Hcp is still secreted in the absence of Tap-1 (Figure 5-2A), suggesting that Tap-1 is essential for secretion of TseL but not for other T6SS proteins. Tap-1 was not detected in culture

supernatants by us and others, suggesting that Tap-1 itself is not secreted (Figure 5-2B and (Altindis et al, 2015)).

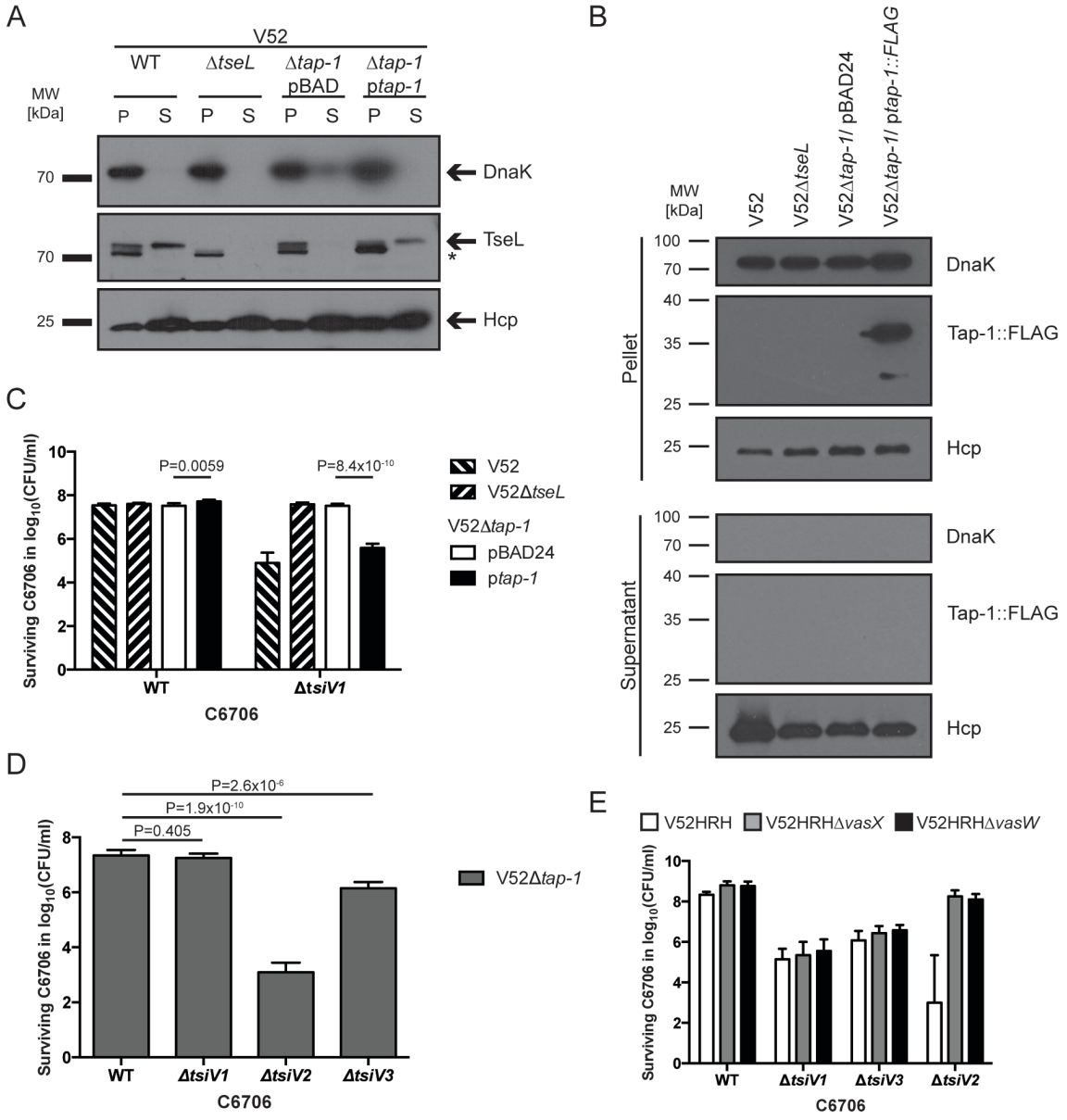


Figure 5-2 Tap-1 is necessary for TseL secretion.

Figure 5-2 Tap-1 is necessary for TseL

(A-B) Tap-1 is required for secretion of TseL but retained in the cytoplasm. Pellet (P) and supernatant (S) of bacterial cultures were analyzed by SDS-PAGE. The result of immunoblotting with antisera against TseL, Hcp and purified antibodies against DnaK and FLAG is shown. The asterisk marks an unspecific band, detected by the antiserum against TseL, that was found in the preimmunized serum. One representative experiment of three independent experiments is shown. (C) Tap-1 is required for TseL-mediated killing. C6706 wild-type or C6706 lacking *tsiV1* were exposed to the indicated V52 mutants in a killing assay. The logarithm of surviving C6706 is shown on the y-axis. The mean ± SD of three independent experiments each performed in duplicate is shown. P-values of a two-tailed, unpaired Student's t-test are indicated. (D) Tap-1 is specifically required for TseL. Killing assays in which V52Δ*tap-1* was mixed with C6706 wild-type and mutants lacking *tsiV1, tsiV2* or *tsiV3*. The logarithm of surviving C6706 is shown on the y-axis. The mean ± SD of three independent experiments each performed in duplicate is shown. P-values of a two-tailed, unpaired Student's t-test are indicated. (E) Killing assay in which V52ΔHRH and indicated mutants were mixed with C6706 wild-type or mutants lacking *tsiV1*, *tsiV2* and *tsiV3,* to assess killing mediated by TseL, VasX and VgrG3, respectively. The mean ± SD from three independent experiments each performed in duplicate is shown.

To determine if the inability of a *tap-1* mutant to secrete TseL prevents killing of other prokaryotic cells in a TseL-dependent manner, we performed a killing assay. V52 or V52Δ*tap-1* was mixed with C6706 or a C6706 mutant lacking the cognate immunity gene *tsiV1*. Under these conditions, C6706 represses its T6SS while maintaining expression of immunity genes (Miyata et al., 2013). The C6706Δ*tsiV1* mutant allows us to analyze TseL-mediated killing because TsiV1 deactivates TseL in the bacterium under attack (here C6706) (Dong et al., 2013; Unterweger et al., 2014). Predator (V52) and prey (C6706) were mixed and plated on nutrient agar plates. After four hours at 37˚C, surviving V52 and C6706 were enumerated. The lack of *tap-1* abolished TseL-mediated killing by V52, comparable to a mutant lacking *tseL* (Figure 5-2C). Complementation with episomal *tap-1* restored TseL-mediated killing. Expression of *tap-1 in trans* in V52 had no effect on the survival of wild-type C6706 but killed the mutant lacking *tsiV1,* indicating that Tap-1 acts on TseL. These results show that TseL-mediated killing depends on Tap-1.

To test whether Tap-1 is required for secretion of effectors in addition to TseL, we determined whether Tap-1 also controls VasX- and VgrG-3-mediated killing. As

indicator strains for VasX- and VgrG3-mediated killing, we used C6706 mutants lacking the cognate immunity gene *tsiV2* and *tsiV3*, respectively. The absence of *tap-1* did not affect killing mediated by VasX or VgrG-3 (Figure 5-2D). These results show that Tap-1 is required exclusively for the effector TseL. An analogous observation was made for VasW, the second protein of the DUF4123 superfamily in *V. cholerae*. VasW is only required for the effector VasX, encoded downstream of *vasW*, but is dispensable for TseL translocation (Figure 5-2E).

Taken together, these data indicate that Tap-1 is required for secretion of TseL, allowing *V. cholerae* to engage in TseL-mediated killing.

5.2.3 VgrG-1 is essential for the secretion of and killing by TseL

The observation that Tap-1 is dedicated to TseL translocation encouraged us to probe other proteins encoded in auxiliary cluster 1 that act in concert with Tap-1. VgrG-1 is encoded immediately upstream of *tap-1* (Figure 5-1A). VgrG-1 in V52 is a structural component of the secretion system and acts as carrier of an effector domain with actin-crosslinking activity (Pukatzki et al., 2007). To determine if VgrG-1 has a dual function (actin crosslinking and TseL translocation), we performed western blot analysis on pellet and supernatant fractions of V52 and a *vgrG-1* mutant. As reported previously (Pukatzki et al., 2007), deletion of *vgrG-1* lowers the amounts of Hcp secreted by cells (Figure 3A). However, TseL secretion was abolished. Providing the *vgrG-1* mutant with episomal *VgrG-1* restored TseL secretion (Figure 5-3A). These observations indicate that VgrG-1 is required for secretion of TseL.

Next, we used our killing assay to determine whether TseL-mediated killing occurred in the absence of VgrG-1. Parental V52 and mutants lacking *tseL* or *vgrG-1* were mixed with C6706 or C6706Δ*tsiV1*. A V52 mutant lacking *vgrG-1* lost its ability for TseL-mediated killing unless provided with episomal VgrG-1 (Figure 5-3B). These results show that *vgrG-1* is required for TseL-mediated killing.

To determine if VgrG-1 is required for translocation of VasX and VgrG-3 in addition to TseL, we performed killing assays using C6706 mutants that lack TsiV2 and TsiV3, the cognate immunity proteins to VasX and VgrG-3, respectively. TseL-mediated

killing was abolished in the absence of *vgrG-1*. VasX- and VgrG-3-mediated killing still occurred in the absence of *vgrG-1* (Figure 5-3C). This finding indicates that VgrG-1 has a dedicated role for TseL translocation.

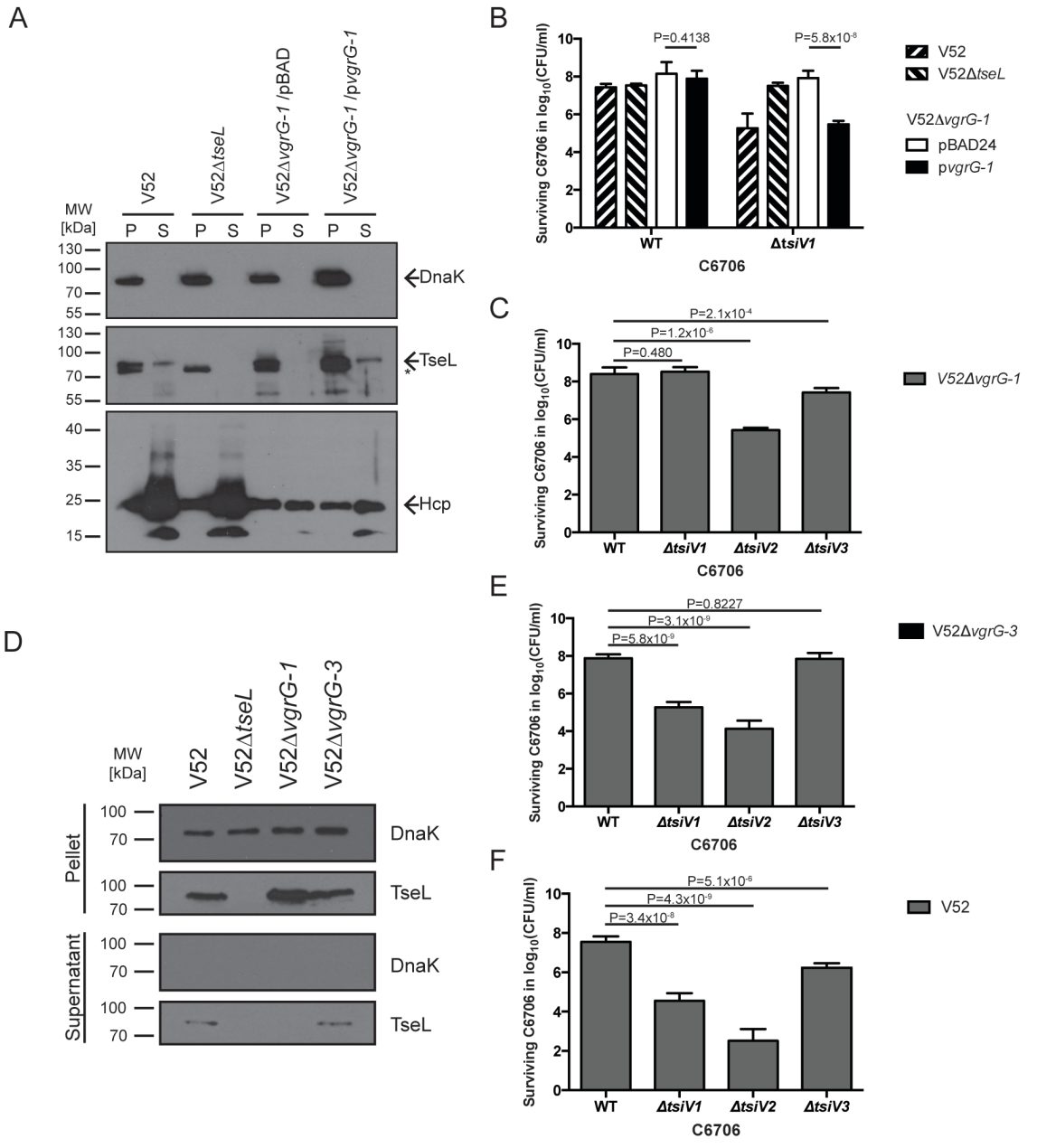


Figure 5-3 VgrG-1 is required for TseL-mediated killing.

Figure legend for Figure 5-3

(A) Analysis of TseL secretion dependence on VgrG-1. Western blot analysis of pellet (P) and supernatant (S) samples of the indicated strains. Samples were immunoblotted with antisera against TseL, Hcp and purified antibodies against DnaK. The unspecific band detected by rabbit serum is marked with an asterisk. One representative of three experiments is shown. (B) Killing assay with wild-type V52 or indicated mutants and wild-type C6706 or a mutant lacking the immunity gene *tsiV1*. The logarithm of surviving C6706 is shown on the y-axis. The mean ± SD of three independent experiments each performed in duplicate is shown. P-values of a two-tailed, unpaired Student's t-test are indicated). (C) V52Δ*vgrG-1* was mixed in a killing assay with wild-type C6706 or mutants lacking immunity genes *tsiV1, tsiV2* or *tsiV3* (specific for TseL, VasX and VgrG-3, respectively). The logarithm of surviving C6706 is shown on the y-axis. The mean ± SD of three independent experiments each performed in duplicate is shown. The P-values of a two-tailed, unpaired Student's t-test are indicated. (D) TseL-secretion of a *vgrG-3* deletion mutant. Pellet and supernatant of indicated bacterial cultures were analyzed by SDS-PAGE. The results of immunoblotting with antisera against TseL and purified antibodies against DnaK are shown. (E) Observation of TseL- and VasX-mediated killing in the absence of *vgrG-3*. The indicated mutants of C6706 were exposed to V52Δ*vgrG-3* in a killing assay to analyze TseL-, VasX- and VgrG-3-mediated killing. The surviving C6706 is shown on the y-axis. The mean ± SD of three independent experiments each performed in duplicate is shown. P-values of a two-tailed, unpaired Student's t-test are indicated. (F) Effector-mediated killing by V52. Killing assay similar to (E). The mean ± SD of three independent experiments each performed in duplicate is shown. P-values of a two-tailed, unpaired Student's t-test are indicated.

To determine if VgrG-3 plays an equally important role for TseL than VgrG-1, we tested the secretion of and killing by TseL in a *vgrG-3*-deficient mutant. First, we analyzed TseL expression and secretion of V52Δ*vgrG-3* in comparison to V52 wild-type or mutants deficient in *tseL* or *vgrG-1*. Similar to previous experiments, TseL was detected in the pellet and supernatant of V52 wild-type but not V52Δ*tseL*. Secreted TseL was not detected in the supernatant of V52Δ*vgrG-1* but in the supernatant of V52Δ*vgrG-3* (Figure 5-3D). DnaK was detected in the pellet only indicating that proteins in the supernatant were actively secreted and not the result of cell lysis. These results suggest that *vgrG-3* is not essential for the secretion of TseL. Second, we tested the ability of V52Δ*vgrG-3* to kill bacteria of the strain C6706 lacking the respective immunity proteins to TseL, VasX and VgrG-3, respectively. Therefore, *vgrG-3*-deficient V52 was mixed with the indicated mutants of C6706 in a killing assay. Killing mediated by TseL and

VasX but not VgrG-3 was observed (Figure 5-3E). For comparison, V52 wild-type is killing via TseL, VasX and VgrG-3 (Figure 5-3F). Taken together, these results indicate that VgrG-3 is not essential for the secretion of and killing by TseL.

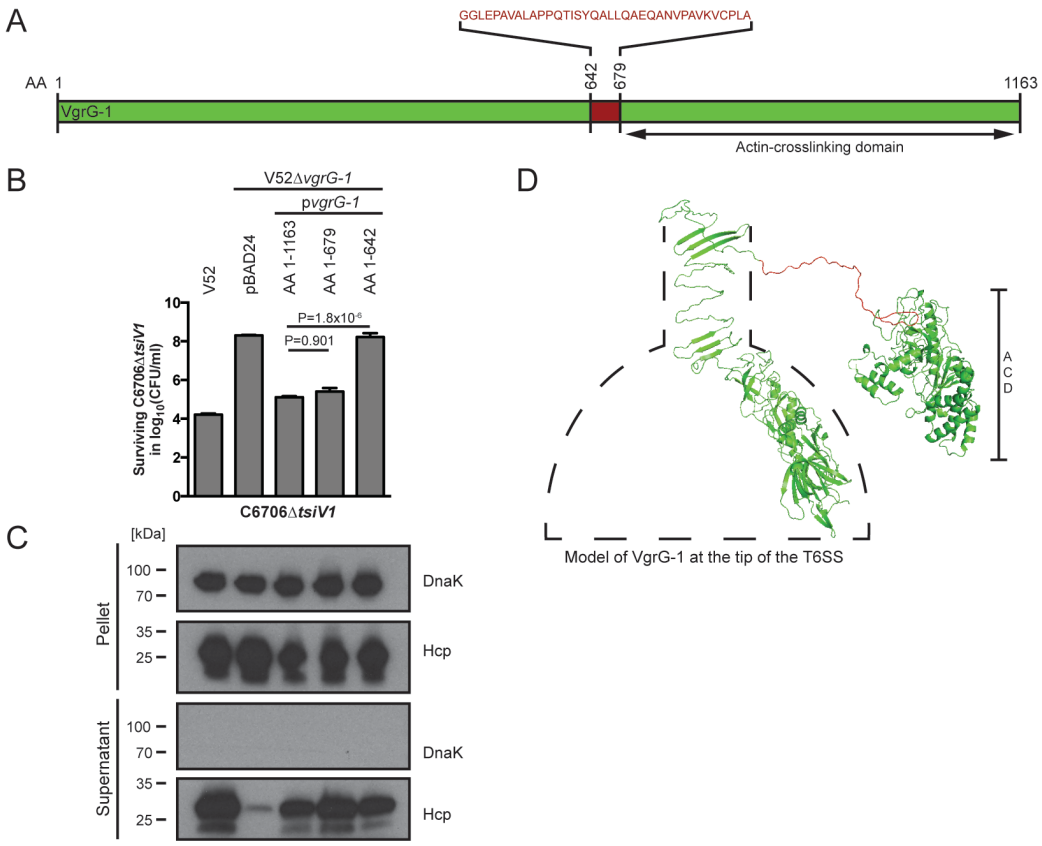


Figure 5-4 Linker-domain of VgrG-1 required for TseL-mediated killing

(A) Graphical depiction of VgrG-1. Amino acid sequence of the highlighted region 642 to 679 is shown. (B) Thirty-seven amino acids of VgrG-1 are sufficient to restore TseL-mediated killing. Killing assay in which a *tsiV1*-deficient mutant of C6706 is exposed to V52 wild-type or V52Δ*vgrG-1* provided with the indicated constructs of *vgrG-1 in trans*. (C) Truncated VgrG-1 versions secrete high levels of Hcp. Pellet and supernatant of V52 strains tested in (B) were analyzed by SDS-PAGE. The results of immunoblotting with antisera against Hcp and purified antibodies against DnaK are shown. (D) Amino acids of interest form the linker region between the core domain and the ACD. Model of VgrG-1 as part of the tip of the T6SS is shown. The ACD is indicated. Amino acids 642-670 are coloured in red.

To identify the region within VgrG-1 that enables TseL-mediated killing, I created constructs that encode amino acids 1-642 and 1-679 of VgrG-1 (Figure 5-4A). I then tested the ability of these construct to mediate killing via TseL in a V52Δ*vgrG-1*

mutant. Whereas the first 679 amino acids of VgrG-1 restored TseL-mediated killing to wild-type levels, the construct encompassing amino acid 1-642 did not (Figure 5-4B). Truncated version of VgrG-1 restored Hcp secretion similar to complementation with full-length VgrG-1 (Figure 5-4C). These 37 essential amino acids from position 642 to position 679 map to a linker region between the core region similar among various VgrG proteins and the C-terminal actin-crosslinking domain unique to VgrG-1 (Figure 5-4D).

In summary, these results indicate that VgrG-1 is essential for secretion of and subsequent killing by TseL.

5.2.4 Tap-1 is required for the interaction between TseL and VgrG-1

The requirements for both VgrG-1 and Tap-1 to secrete TseL led us to investigate how these proteins contribute to the secretion process of TseL. VgrG-1 is secreted with other VgrG proteins (Pukatzki et al., 2007) and PAAR-domain-containing proteins (Shneider et al., 2013) as part of a macromolecular complex at the tip of the secretion system, thus we hypothesized that TseL interacts with Tap-1 and VgrG-1 for recruitment to this complex. To test physical interactions among Tap-1, VgrG-1 and TseL, we performed immunoprecipitation experiments followed by western blot analysis. First, we prepared lysates from V52Δ*tap-1* cells transformed with empty vector, p*tap-1*::His, or p*tap-1*::FLAG. FLAG-tagged Tap-1 was precipitated with anti-FLAG-M2 affinity beads. His-tagged Tap-1 is not expected to bind to the anti-FLAG-M2 affinity gel and allowed us to test for unspecific binding of TseL to the beads. Lysates and immunoprecipitates were subjected to SDS PAGE, and analyzed for TseL, Tap-1::His and Tap-1::FLAG. TseL was found in the lysates of all three strains, but only in the immunoprecipitate generated with FLAG epitope (Figure 5-5A). His-tagged Tap-1 was found in the pellet but not in the supernatant. No band was detected with anti-TseL, anti-His or anti-FLAG antibodies in precipitated lysates of V52Δ*tap-1* transformed with empty vector. This analysis established an interaction between Tap-1 and TseL.

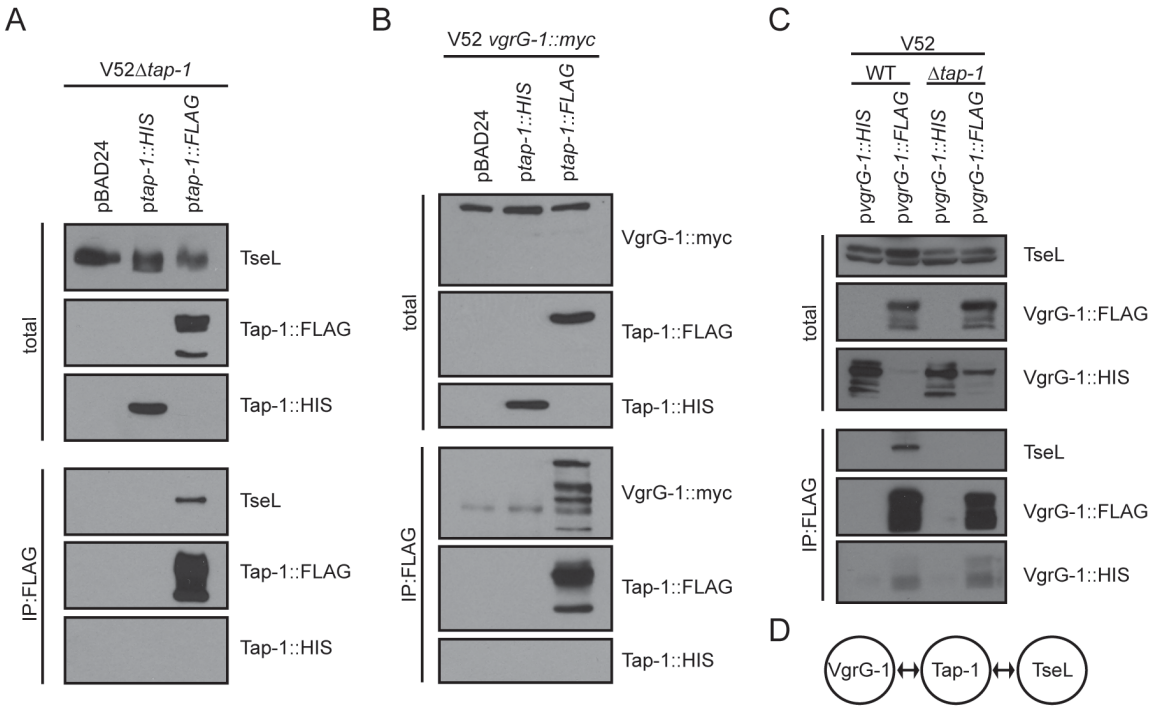


Figure 5-5 Tap-1 mediates interaction between TseL and VgrG-1.

(A and B) Interactions between Tap-1 and TseL (A) or VgrG-1 (B). Shown are immunoblots of lysates (total) and immunoprecipitates with anti-FLAG affinity beads (IP:FLAG) of V52Δ*tap-1* (A) or V52 *vgrG-1::myc* (B) transformed with empty vector or a plasmid encoding His-tagged or FLAG-tagged Tap-1. (C) Interaction between TseL and VgrG-1. Shown are immunoblots of lysates (total) and immunoprecipitates with an anti-FLAG affinity gel (IP:FLAG) of V52 or V52Δ*tap-1* transformed with empty vector or a plasmid encoding HIS-tagged or FLAG-tagged VgrG-1. (D) Diagrammatic interpretation of panels A–C. Arrows indicate interaction.

To test if Tap-1 also interacts with VgrG-1, we analyzed lysates of V52 with a myc-tagged version of VgrG-1 that was transformed with either empty vector, p*tap-1*::His or p*tap-1*::FLAG. VgrG-1 was found in the lysates of all three strains but only in the immunoprecipitate with Tap-1::FLAG (Figure 5-5B). No band with the size of VgrG-1 was detected in the immunoprecipitate of lysates that contained His-tagged Tap-1 or untagged Tap-1. This suggests that VgrG-1 interacts with Tap-1.

As Tap-1 interacts with both VgrG-1 and TseL, we wanted to test the role of Tap-1 in the interaction between VgrG-1 and TseL. We tested lysates of V52 or V52Δ*tap-1* that express either His-tagged or FLAG-tagged VgrG-1. Anti-FLAG affinity beads were added to the lysates, pulled down and analyzed with antibodies against TseL, FLAG or

His. TseL was detected in the immunoprecipitates with VgrG-1::FLAG only in V52 wild-type, not in the absence of Tap-1 (Figure 5-5C).

These results suggest that TseL has the ability to interact with Tap-1, which can interact with VgrG-1 (Figure 5-5D). TseL depends on Tap-1 to form an interaction with VgrG-1, because we were unable to detect TseL in the immunoprecipitate of FLAG-tagged VgrG-1 from a lysate lacking Tap-1. We also observed that Tap-1 has the ability to form homo-multimers (Appendix A, Figure 9-7).

5.2.5 Diversity of Tap-1 among *V. cholerae* strains

We previously showed that *V. cholerae* strains encode different sets of T6SS effector modules within their gene clusters (Unterweger et al., 2014). Auxiliary cluster 1 encodes one of two effectors depending on the strain (Unterweger et al., 2014). TseL is encoded in the T6SS gene clusters of strains like the pandemic O1 serogroup strain N16961 and belongs to the A effector family. Effectors of the B family share less than 30% amino acid sequence identity with TseL and are encoded in the T6SS gene clusters of strains like the O39 serogroup strain AM-19226 or the O12 serogroup strain 1587. Having established a role for VgrG-1 and Tap-1 in the secretion of TseL (Figure 5-2 and Figure 5-3), we hypothesized that VgrG1 and Tap-1 have to be tailored for the translocation of A or B family effectors. First, we investigated the differences in VgrG-1 of different strains. We compared the amino acid sequences of VgrGs encoded by *V. cholerae* strains used previously for our analysis of effector modules (Unterweger et al., 2014). Twenty-five out of thirty-six strains encode a VgrG-1 protein that consists of the core domains and the C-terminal extension with actin crosslinking activity (Figure 5-6). V52 and AM-19226 are examples of strains that contain a VgrG-1 with an actin crosslinking domain (ACD) (Figure 5-7A). 1587 is an example of a strain that contains a VgrG-1 without an ACD (Figure 5-7A). This analysis shows that VgrG-1 proteins of *V. cholerae* strains differ from each other by the presence or absence of a C-terminal extension.

Strain	VgrG-1 ACD	Tap-1 clade	Effector family
V51	-	3	B
1587	-	3	B
DL4211	+	2	B
VL426	+	2	B
MZO-2	+	1	A
12129(1)	-	3	B
623-39	-	3	B
AM-19226	+	2	B
HE-25	+	2	B
HE-39	-	3	B
HE-45	-	3	B
HC-43B1	-	3	B
Amazonia	-	3	B
TM11079-80	-	3	B
DL4215	+	2	B
TMA21	-	3	B
MZO-3	-	3	B
2740-80	+	1	A
M66-2	+	1	A
BX330286	+	1	A
MAK757	+	1	A
CIRS101	+	1	A
CP1041	+	1	A
HC-07A1	+	1	A
HC-32A1	+	1	A
HC-33A2	+	1	A
HC-38A1	+	1	A
MO10	+	1	A
N16961	+	1	A
2010EL-1786	+	1	A
C6706	+	1	A
MJ1236	+	1	A
V52	+	1	A
CA401	+	1	A
O395	+	1	A
NIH41	+	1	A

Figure 5-6 VgrG-1, Tap-1 and effectors in various *V. cholerae* strains.

For each of the 36 *V. cholerae* strains analyzed, we show the presence (+) or absence (-) of the actin crosslinking domain (ACD) in VgrG-1, the clustering of Tap-1 to clade 1, 2 or 3, and the family of the effector (A or B) encoded in the auxiliary cluster 1.

Next, to analyze the diversity of Tap-1 proteins among *V. cholerae* strains, we aligned the amino acid sequences of Tap-1 from the same 36 strains. Phylogenetic analysis showed that Tap-1 clusters into three clades Figure 5-7B). An expanded analysis of all sequenced *V. cholerae* strains available in the NCBI database did not reveal additional Tap-1 alleles. Tap-1 from strains V52, AM-19226 and 1587 exemplifies clades 1, 2 and 3, respectively (Figure 5-7B). Alignment of three representative proteins from each cluster show that they share multiple 15- to 33-residue-long motifs in the N-terminal segment (grey regions in Figure 5-7C). Motifs that make up domain DUF4123 are found in these highly conserved regions, except for the first motif, which is found in a non-conserved region in the polymorphic N-termini and varies at two amino acid positions between the analyzed sequences (Figure 5-7C). Differences in three

representative sequences of the N-terminal segment (yellow and orange) and differences in the C-terminal segment (in blue and purple) are indicated in Figure 5-7C. Sequences in cluster 2 (e.g. Tap-1^{AM19226}) and 3 (e.g. Tap-1^{1587}) share highly similar C-terminal segments but varying N-terminal segments (Figure 5-7C). Sequences in cluster 1 (e.g. Tap-1^{V52}) contain highly similar N-terminal segments compared to sequences in cluster 2 (e.g. Tap-1^{AM19226}) but contain distinct C-terminal segments that share less than 30% amino acid identity (Figure 5-7C). These analyses demonstrate that different yet distinct forms of Tap-1 are distributed among *V. cholerae* strains.

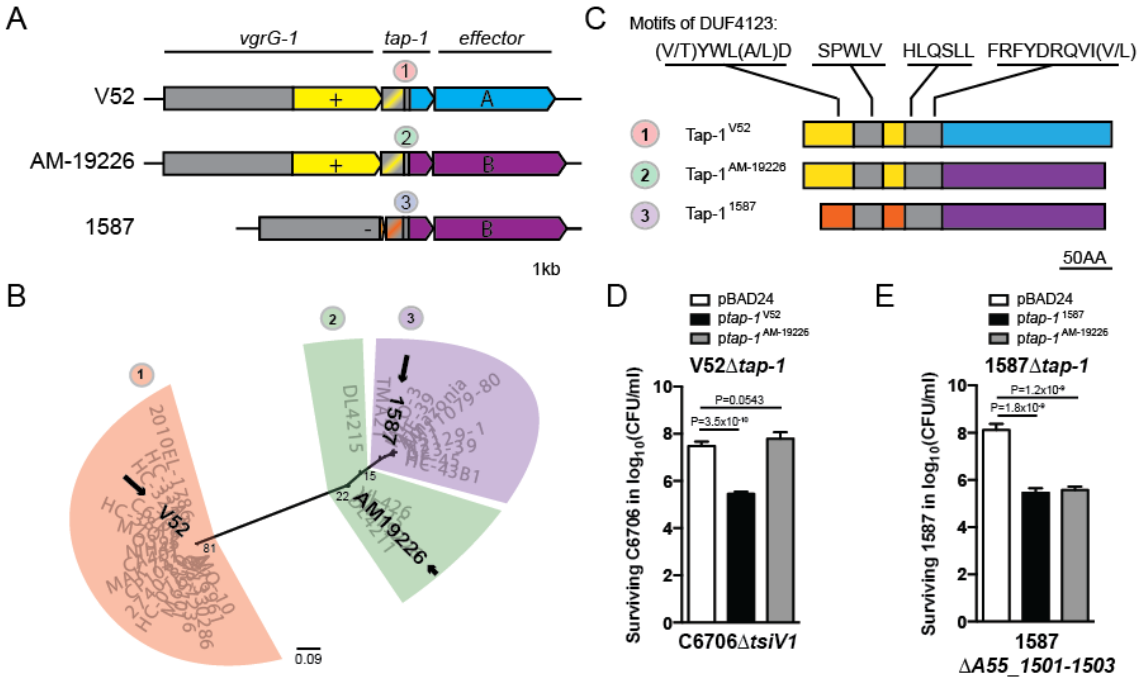


Figure 5-7 Diversity of Tap-1 among *V. cholerae* strains.

Figure 5-7 Diversity of Tap-1 among *V. cholerae* strains
(A) Cartoon of *vgrG-1*, *tap-1* and the effector protein-encoding gene in the auxiliary cluster 1 of three indicated *V. cholerae* strains. Same colour indicates more than 90% identity of the amino acid sequences encoded in the genes compared. Different colour indicates less than 30% identity and mixed colour indicates 30–90% identity. The presence or absence of the ACD encoded in *vgrG-1* is indicated with + and -, respectively.(B) Diversity of Tap-1. Phylogenetic analysis of Tap-1 encoded in 36 *V. cholerae* strains. Analysis was performed using Raxml (Stamatakis, 2014). The names of the strain harbouring the Tap-1 are indicated in the tree. Three clades are highlighted with different colours. Three representative sequences that cluster to different clades are highlighted with an arrow and used for analysis. (C) Cartoon of Tap-1 from the three different clades. Regions that share more than 80% identity in the three amino acid sequences are shown in the same colour, regions with less than 45% identity are shown in different colours. The conserved sequences of the motifs of DUF4123 and their positions are indicated. (D) Specificity of C-terminal segment of Tap-1. Killing assay in which the lack of *tap-1* in V52 is complemented by empty vector (control) or one of two different alleles of *tap-1*. The ability to kill in a TseL-dependent manner is indicated by killing or survival of C6706Δ*tsiV1*. The arithmetic mean ± SD of log-transformed data of three independent experiments, each performed in duplicate, is shown. P-values of a two-tailed, unpaired t-test are indicated. (E) Specificity of N-terminal segment of Tap-1. Killing assay in which the lack of *tap-1* in 1587 is complemented by empty vector (control) or one of two different alleles of *tap-1*. The ability to facilitate killing by the effector A55_1502 is indicated by killing or survival of 1587ΔA55_1501-03. The arithmetic mean ± SD of log-transformed data of three independent experiments, each performed in duplicate, is shown. P-values of a two-tailed, unpaired t-test are indicated).

To test how differences in the N- or C-terminal segments of Tap-1 affect the function of Tap-1, we compared the adaptor proteins in pairwise combinations (see Figure 5-8 for extended analysis of Tap-1 alleles in different strain backgrounds). The adaptors from V52 and AM-19226 share similar N-terminal segments but differ in their C-terminal segments (Figure 5-7 C). To determine the effector specificity of the C-terminal portion of Tap-1, V52Δ*tap-1* was provided with episomal *tap-1*^{V52} or *tap-1*$^{AM-19226}$ (Figure 5-7 D). All Tap-1 alleles were expressed (Figure 5-8 B), but only Tap-1^{V52} facilitated TseL-mediated killing of C6706Δ*tsiV1* at levels similar to wild type; Tap-1$^{AM-19226}$ did not support killing (Figure 5-7 D). This indicated that differences in the Tap-1 C-terminal segment are not tolerated.

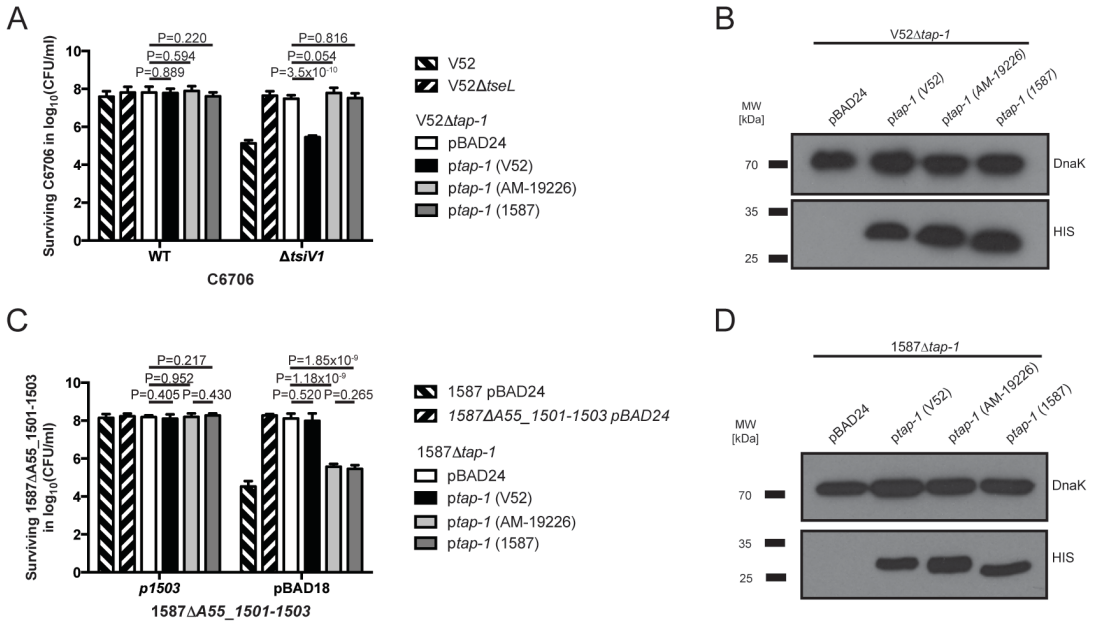


Figure 5-8 Specificity of Tap-1 in various strain backgrounds.

(A) Killing assay in which the lack of *tap-1* in V52 is complemented by empty vector (control) or one of three different C-terminally HIS-tagged alleles of *tap-1*. The ability to kill in a TseL-dependent manner is indicated by killing or survival of C6706Δ*tsiV1*. The arithmetic mean ± SD of log-transformed data from three independent experiments, each performed in duplicate, is shown. P-values of a two-tailed, unpaired Student's t-test are indicated. (B) Analysis of the expression of the HIS-tagged Tap-1 alleles by SDS-PAGE. Samples were subject to western blotting with anti-DnaK and anti-HIS antibodies. (C) Killing assay as in (A) except using 1587Δ*tap-1* as predator and 1587Δ*A55_1501-03* transformed with pBAD18 as prey. 1587Δ*A55_1501-03* lacks the genes encoding the adaptor, effector and immunity protein of auxiliary cluster-1. (D) Western blot analysis as described in (B) of the indicated samples.

To test whether differences in the N-terminal segment are tolerated, we compared the adaptor proteins from 1587 and AM-19226, which share common C-terminal segments. Both also share the conserved motives of the DUF4123 domain, but differ in the remaining sequence of the N-terminal segments. A 1587Δ*tap-1* mutant was provided with episomal *tap-1*$^{AM\text{-}192262}$ or *tap-1*^{1587}. All Tap-1 alleles were expressed (Figure 5-8D). Both adaptor proteins restored effector-mediated killing (Figure 5-7E). This indicated that differences in the N-terminal segment are tolerated in the strain background tested. We were unable to test the Tap-1 alleles in the AM-19226 background because no effector-mediated killing could be detected in that strain. This could result from orphan immunity genes or from lack of expression of the effector module in AM-19226.

Taken together, these data show that *V. cholerae* strains encode a variety of VgrG-1 and Tap-1 proteins. The VgrG-1 proteins differ from each other by the presence or absence of the enzymatically active C-terminal extension. The Tap-1 proteins differ from each other in their N- and C-terminal segments, with the C-terminal segments being specific for the T6SS effector. We found that the N-terminal segment of Tap-1 correlates with the presence or absence of a C-terminal extension of VgrG-1; however, both N-terminal Tap-1 segments appear to be operational in strains that miss the ACD domain of VgrG-1.

5.2.6 N-terminus of Tap-1 evolved under diversifying selection

The N-terminal segment of Tap-1 correlates with the presence or absence of a C-terminal extension of VgrG-1 (Figure 5-7). Despite the polymorphic nature of the N-terminal segment of Tap-1, the N-terminal segments of the 36 analyzed sequences still share an overall amino acid sequence identity of 64% (Figure 5-9A), indicating that they are homologs (Pearson, 2013).

We hypothesized that the two different N-terminal segments evolved from the same common ancestral segment. During this process, different amino acid sequences might have been selected for. Therefore, we analyzed retention of amino acid sequence versus nucleotide sequence by studying the effects of single nucleotide polymorphisms on the amino acid sequences of Tap-1 from all 36 analyzed strains (Figure 5-7). As an indicator for the number of non-synonymous mutations (change of amino acid sequence) and synonymous mutations (no change of amino acid sequence), we determined the ratio of percentage amino acid identity to nucleotide sequence identity. We subdivided the 5' end and N-terminal sequence into three regions, plus the central region. The short, terminal 14-residue long region of the N-terminal segment unique to class 1 and 2 Tap-1 proteins (Figure 5-7) was excluded from this analysis, because this region would bias the analysis when comparing sequences of class 1 and 2 to class 3 Tap-1 proteins lacking this region. Functional analysis demonstrated that this short region is not required for a functional Tap-1, thus supporting our decision to exclude this fragment from our analysis. We observed that region 2 has a ratio of less than one (Figure 5-9B), demonstrating that

the relative number of nucleotide substitutions is higher than the relative number of amino acid changes. This indicates selection pressure to retain the amino acid sequence but not the nucleotide sequence in region 2 (Suzuki and Gojobori, 1999).

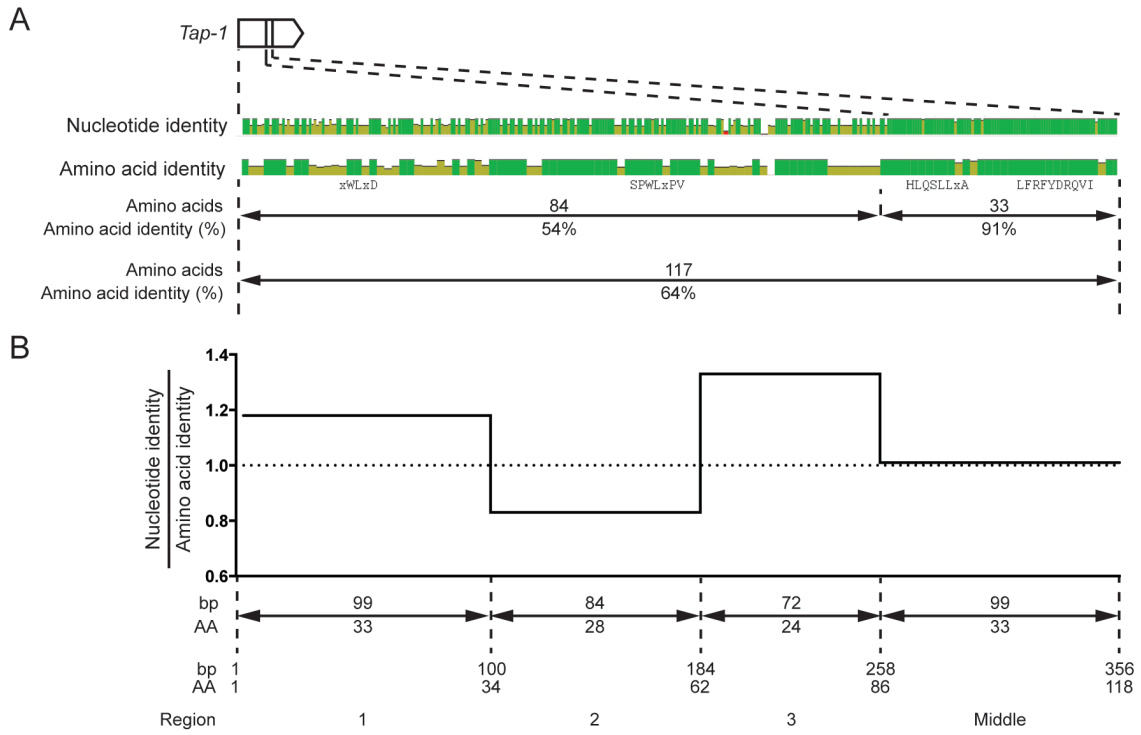


Figure 5-9 N-terminal domain of Tap-1 is under diversifying evolutionary selection.

(A) N-termini of Tap-1 share more than 30% overall amino acid sequence identity. The identity of the nucleotide and amino acid sequences over the indicated regions among *tap-1* of 36 *V. cholerae* strains (sliding window=1) is shown. Green bars indicate 100% identity, yellow bars 30–99% and red bars less than 30%. The percent amino acid identity over the indicated region, with the length of amino acids indicated, is shown at the bottom of the figure. The first 14 amino acids of class 1 and 2 Tap-1 proteins were excluded from this analysis. (B) The ratio of percentage nucleotide sequence change to percentage amino acid sequence change over the indicated region is shown in the graph.

Regions 1 and 3 have a ratio greater than one (Figure 5-9B). Relatively few nucleotide changes cause a relatively high number of amino acid changes, indicating that these regions are selected to diversify.

Motifs of DUF4123 are found in regions 1, 2 and the middle region, and are highly conserved on the amino acid level (Figure 5-9B). Even the WLxD motif found in the diverse region 1 is conserved, further highlighting the importance of this domain.

In summary, regions in the N-terminal segment of Tap-1 show signs of selection to conserve specific regions and to diversify others. The region in the middle of Tap-1 is highly conserved on both the nucleotide and amino acid levels, indicating its importance on both levels.

5.2.7 Acquisition of ACD affects effector-mediated killing

The analysis of the N-terminus of Tap-1 suggests diversifying selection that could be imagined in response to a selection pressure that affects effector-mediated killing. I hypothesize that the acquisition of the ACD in strain 1587 affects effector mediated killing.

To test this hypothesis, I inserted the ACD into the chromosome of 1587 (Figure 5-10A). The resulting strain was tested for T6SS-mediated killing by the effector encoded downstream of *vgrG-1* in the same cluster in a killing assay. Widl-type 1587 or the insertion mutant were mixed with 1587Δadaptor, effector, immunity (1587Δaei) that either contains empty vector or expresses the immunity protein-encoding gene (A55_1503) in trans. Equal amounts of around 10^7 survivors of 1587Δaei p1503 were recovered independent of the exposure of this mutant to wild-type 1587 or the ACD insertion mutant of 1587 (Figure 5-10B). When the mutant 1587Δaei harbouring pBAD24 alone was exposed to wild-type 1587, 1000 fold effector-mediated killing was observed. When the mutant 1587Δaei pBAD24 was exposed to the ACD insertion mutant of 1587, 100 fold killing was observed. This 10 fold difference between the two indicates that the ACD insertion into the chromosome affects but does not abolish T6SS-mediated killing and thus provides a potential selection pressure to regain full amounts of killing.

Further experiments are needed to determine if the ACD affects the activity of this effector or T6SS activity in general.To test if a modification of the adaptor protein can restore full effector activity, I suggest to generate a deletion mutant of the adaptor protein-encoding gene in the 1587 *vgrG-1::ACD* background. This deletion mutant could

be complemented with an adaptor that contains the N-terminus specific for the ACD and analyzed in a killing assay for effector-mediated killing.

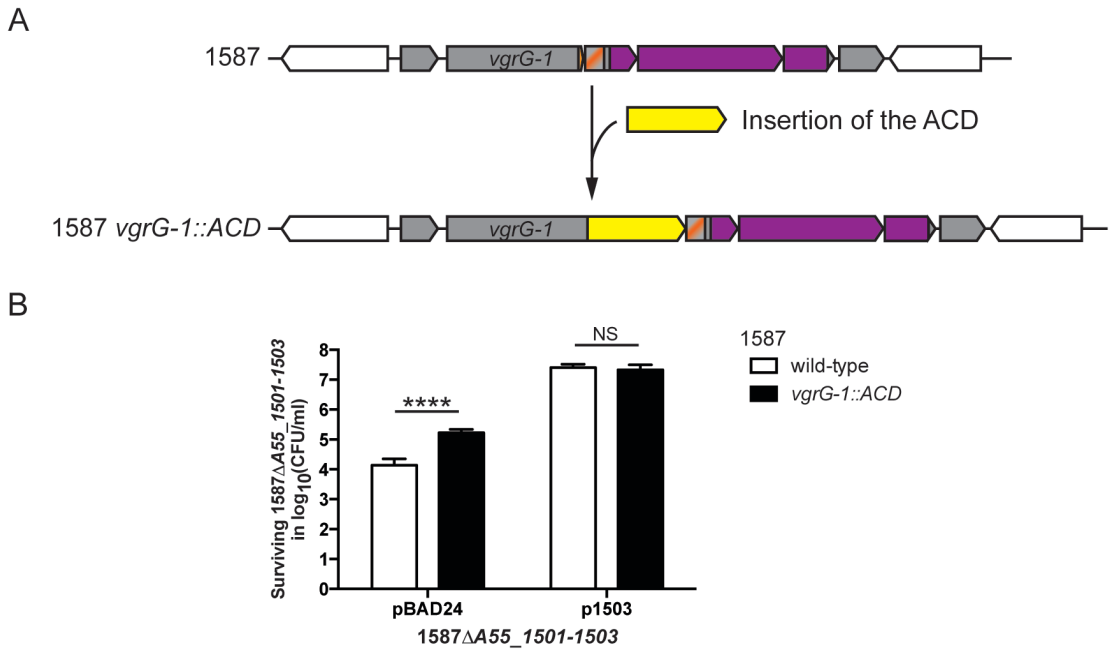


Figure 5-10 Gain of the ACD affects effector-mediated killing

(A) Graphical depiction of the auxiliary cluster 1 before and after the insertion of the actin-crosslinking domain. (B) Killing assay testing the ability of 1587 with or without the ACD to kill in dependency on the effector encoded in auxiliary cluster 1. The arithmetic mean ± SD of log-transformed data of three independent experiments, each performed in duplicate, is shown. The results of a two-tailed, unpaired t-test are indicated (**** P<0.0001; NS P>0.05).

5.2.8 The middle region of *tap-1* is conserved at the nucleotide level

Genes encoding T6SS effectors and immunity proteins differ in their GC content from the surrounding genes, indicating their acquisition and exchange among *V. cholerae* strains by horizontal gene transfer (Unterweger et al., 2014). For example, the strain VL426 is closely related to MZO-2 but encodes an effector other than *tseL* in its auxiliary cluster 1 (Unterweger et al., 2014). In contrast to the N-terminal segments of Tap-1 in our 36 analyzed sequences, the C-terminal segments of Tap-1 share less than 30% overall amino acid identity (Figure 5-11A). This indicates a lack of common ancestry for the C-

terminal segments and suggests different evolutionary origins. We hypothesized that *tap-1* contains a recombination site to create a diverse Tap-1 family.

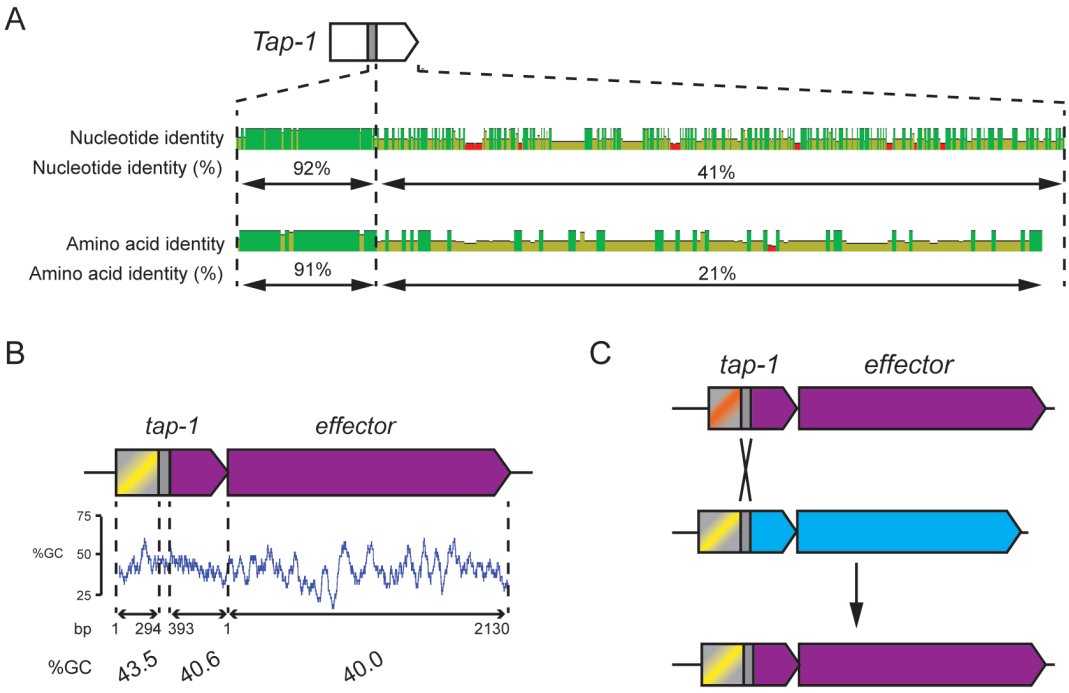


Figure 5-11 *Tap-1* encodes a putative recombination site.

(A) Alignments of the nucleotide and amino acid sequences of the Tap-1 C-termini. Identity in 36 amino acid sequences of Tap-1 or nucleotide sequences of *tap-1* are shown (sliding window=1). Green bars indicate 100% identity, yellow bars indicate 30–99% and red bars indicate less than 30%. Average percent identity is shown over the regions indicated with arrows. (B) Analysis of the GC content (sliding window=50) of *tap-1* and the effector-encoding gene in *V. cholerae* strain 1587. Average percent GC in regions indicated with arrows is shown. (C) Model for the acquisition of a new 3' end for *tap-1* and an effector-encoding gene as a result of the recombination site within *tap-1*.

Analysis of the nucleotide sequences of 36 *tap-1* genes revealed a highly conserved 99 bp region in the middle of the gene, with 92% identity (Figure 5-11A). Analysis of GC content shows that recombination likely happened at this site because the GC content differs by over 3% upstream and downstream of this site (Figure 5-11B). The 3' end of *tap-1* differs in its GC content from the effector-encoding gene downstream by only 0.6%, indicating that the 3' end of *tap-1* plus the effector-encoding gene downstream could have been acquired and exchanged together between strains. Thus, we propose that

the 3' end of *tap-1* co-segregates with the effector-encoding gene downstream. Recombination in the core region would explain the lack of common ancestry among diverse C-terminal segments of Tap-1. Recombination within *tap-1* would maintain the open reading frame and give rise to a new protein with a different function than the proteins from which it derived its component pieces (Figure 5-7), thus we call Tap-1 a chimera.

5.3 Discussion

The competitive fitness of *V. cholerae* strains in regard to dealing with competitors, including other *V. cholerae* strains, is determined by the set of encoded T6SS effector modules (Unterweger et al., 2012; Unterweger et al., 2014). In this study, we took bioinformatics, genetic and biochemical approaches to identify and characterize T6SS components required for the translocation of diverse effector proteins encoded in one of the three T6SS effector modules, namely auxiliary cluster 1. Our analysis of the newly characterized adaptor protein Tap-1 revealed a chimeric composition with high diversity, which may have evolved through adaptation and recombination. The proposed mechanism of Tap-1-dependent effector translocation might especially be important for diverse, newly acquired effector proteins, which rely on a conserved T6SS for translocation. These findings advance our understanding of how VgrG proteins and the adaptor protein Tap-1 mediate T6SS effector translocation. Considering the numerous types of bacteria with a T6SS secretion system and the ubiquity of proteins belonging to the DUF4123 superfamily, the implications of our findings are likely to reach well beyond *V. cholerae*.

We propose a model in which the C-terminal segment of Tap-1 binds the cognate downstream-encoded effector TseL. The N-terminal segment of Tap-1 then binds VgrG-1, which recruits the TseL effector to the tip of the T6SS central Hcp tube in preparation for translocation into a neighboring cell. As we do not detect Tap-1 in the supernatant of *V. cholerae* cells engaged in constitutive T6SS-mediated translocation of TseL, we hypothesize that Tap-1 hands over TseL to VgrG-1 prior to ejection of the effector-loaded central Hcp tube. Such a mechanism would retain Tap-1 in the cytoplasm as

shown in Figure 5-2, making Tap-1 available for loading of the next Hcp tube. A recent study by Whitney and colleagues on the T6SS of *Pseudomonas aeruginosa* and *Edwardsella cloacae* showed that VgrG proteins are also required for translocation of PAAR-domain containing effectors (Whitney et al., 2014) (Shneider et al., 2013). We speculate that TseL interacts with VgrG at a different site than PAAR-domain effectors, because TseL is missing the PAAR-domain. To which extent the herein described interactions are direct or depend on additional proteins needs to be further elucidated. A direct interaction between VgrG-3 and TseL has previously been proposed based on the detection of TseL in immunoprecipitates of VgrG-3 (Dong et al., 2013). Our results do not exclude such an interaction, but rather suggest that VgrG-1 is indispensable for TseL whereas VgrG-3 is dispensable.

Our characterization of Tap-1 reveals that a protein of the DUF4123 superfamily is required for the interaction between a T6SS effector and its cognate VgrG protein, subsequent effector secretion, and effector-mediated killing. We made similar observations for another member of the DUF4123 superfamily, VasW (Miyata et al., 2013). Genes encoding DUF4123 domains are often located upstream of effector genes with a MIX (marker for type VI effectors) motif (Salomon et al., 2014), suggesting that one conserved role for DUF4123-containing proteins might be to assist in T6SS-mediated effector delivery. The function of Tap-1 is characteristic of a chaperone protein in terms of binding cargo effectors and delivering them to the secretion apparatus as described for chaperons in a variety of secretion systems, including the type III and type VII secretion system (Daleke et al., 2012; Page and Parsot, 2002). Such a chaperone function has been described for Hcp from *P. aeruginosa*, which harbours T6SS effectors smaller than 20 kDa (such as Tse2) inside the Hcp conduit prior to ejection (Silverman et al., 2013). In the absence of Hcp, Tse2 becomes unstable. Proteins of the DUF4123 superfamily might also stabilize their cognate effector, as a recent study by Ma and colleagues suggests. When the authors purified recombinant Tde1 an *Agrobacterium tumefaciens* T6SS effector, higher yields were obtained when co-expressed with a member of the DUF4123 protein superfamily (Ma et al., 2014b). Tap-1 may also have a stabilizing function for TseL that we may have not detected due to the non-quantitative nature of luminol-based chemiluminescent detection. In summary, these findings show that certain T6SS effectors

require additional proteins for stability or secretion, and that different types of effectors depend on different types of accessory proteins like Tap-1 or Hcp.

Our analysis of VgrG-1 reveals that not all VgrG-1 proteins among *V. cholerae* strains contain a C-terminal extension that confers anti-eukaryotic activity. The actin crosslinking domain (ACD) of VgrG-1 is an example of an effector domain that is missing in some strains, not replaced by another effector domain. The T6SS gene clusters of *V. cholerae* thus not only differ between strains in their composition of T6SS effectors but also in their total number of effectors.

We observed that *V. cholerae* strains encode a diversity of Tap-1 proteins that differ in their N- and C-terminal segments. The N-terminal segment of Tap-1 of *V. cholerae* strains that contain VgrG-1 with an ACD differs from the N-terminal segment of Tap-1 of strains that contain VgrG-1 without an ACD. The C-terminal segments differ between Tap-1 of strains that differ in their anti-prokaryotic effector, thus the C-terminal segment appears to display specificity for the cognate effector they encode. Of the four possible combinations of the two types of N-terminal and C-terminal segments, we found only three combinations in an exhaustive search of publically available genomic sequences for *V. cholerae* strains. The lack of a Tap-1 chimera with the N-terminal segment found in strains without the ACD and the C-terminal segment specific for TseL suggests a functional relationship between ACD and TseL.

Key to the chimeric structure of Tap-1 is the putative recombination site in the middle of *tap-1*. We observed that the 3' end of *tap-1* downstream of the recombination site co-segregates with the effector-encoding gene. We now include this 3' end of *tap-1* in our definition of a T6SS effector module, in addition to genes encoding an effector protein and an immunity protein. We envision a crucial role for *tap-1* during the exchange of effector modules that determines compatibility. Our previous phylogenetic analysis of *V. cholerae* strains found that distantly related strains encode the same effector module set (Unterweger et al., 2014). This could be explained by exchange of effector modules through horizontal gene transfer and subsequent compatibility group switching. Despite this possibility of T6SS module exchange, pandemic strains all encode the same module set, AAA (Unterweger et al., 2014). The fact that pandemic strains have conserved T6SS effectors indicates selection for their module set. All herein analyzed

pandemic strains also contain a VgrG-1 with an ACD. Thus, it appears that the ACD and the AAA module set comprise the characteristic effector module set of pandemic strains in addition to the presence of cholera toxin and toxin co-regulated pilus.

T6SS-mediated competition interferes with cell-cell contact and thereby affects contact-dependent exchange of genetic elements. When *V. cholerae* inhabits chitin surfaces of copepods in environmental reservoirs, the T6SS and competence (the ability to take up extracellular DNA) are coupled (Borgeaud et al., 2015). Our *compatibility rule* predicts that a predator *V. cholerae* strain will take up the genomic DNA from incompatible prey that succumb to the T6SS-mediated attack and are lysed. Any double crossovers that include a complete T6SS cluster like the small auxiliary cluster containing *tseL* will replace the recipient host cluster with the donor prey cluster. This prey cluster should be functional in the host cell, because the prey cluster provides the proper adaptor gene that operates with the new effector. However, if the recipient carries a VgrG-1 with an actin-crosslinking domain and the prey does not, then the recipient gains a new activity, but loses the ability to crosslink actin. We hypothesize that recombination in the *tap-1* core splices a new 3' end to the existing 5' end of *tap-1* on the chromosome, allowing the cell to use its resident T6SS to secrete newly acquired effectors, while retaining the ACD on VgrG-1. In consequence, the recipient predator bacterium would change its compatibility group, allowing the strain to kill bacteria with which it was previously compatible - including its parental kin. The ability to horizontally acquire genes from incompatible strains in the environment permits strains to diversify and to select for adaptive traits beneficial for persistence in the environment or the host.

Chapter 6

Classical *V. cholerae* strains are weak microbial competitors

Stefan Pukatzki recovered the T6SS gene cluster of PA1849. Ashley Wilton generated data for Figure 6-3B and Figure 6-9.

6 Classical *V. cholerae* strains are weak microbial competitors

6.1 Introduction

The current 7th cholera pandemic follows six cholera pandemics of the 19th and 20th centuries (Pollitzer, 1959b). The 7th pandemic is caused by O1 serogroup strains of the El Tor biotype, whereas the six pandemics leading up to the 7th pandemic were caused by O1 serogroup strains of the classical biotype (Faruque et al., 1998). How strains of the classical biotype got replaced by strains of the El Tor biotype in causing pandemics remains a mystery.

Strains of the classical biotype differ from strains of the El Tor biotype in their phenotype, their genotype, and their virulence potential (Chun et al., 2009; Kaper et al., 1995). Historically, the two biotypes are differentiated from each other based on the ability of El Tor strains to agglutinate chicken erythrocytes and to cause a colour change indicating acetoin production in the Voges-Proskauer reaction (Kaper et al., 1995). In contrast to El Tor biotype strains, classical strains can be inhibited by polymyxin B (Kaper et al., 1995) and lysed by biotype-specific phages like the classical IV bacteriophage (Mukerjee, 1963) and the FK bacteriophage (Takeya et al., 1981). O1 serogroup strains are further classified based on their surface antigens into the serotypes Ogawa and Inaba (Sakazaki and Tamura, 1971). Analysis of the whole genome sequences of strains of classical and El Tor biotypes segregates the biotypes into the two different phylocore genome subclades 1 and 2, respectively (Chun et al., 2009; Devault et al., 2014). The phylocore genome clade of pandemic strains is defined by the gene content shared by all pandemic strains (Chun et al., 2009). By acquiring genetic elements and islands, a common ancestor of all pandemic strains might have evolved into two different lineages that are now called phylocore genome subclades 1 and 2 (Chun et al., 2009). Strains of the two biotypes contain different alleles of *ctxB* which encodes the cholera toxin B subunit. Classical strains encode two copies of the *ctx* operon (Mekalanos, 1983; Moseley and Falkow, 1980) in their genome whereas most El Tor strains encode a single copy of the *ctx* operon. An example of an El Tor strain with two *ctx* copies is HK1, which was isolated in Hong Kong in 1961 (Mekalanos, 1983). Strains of the classical biotype cause more severe disease symptoms than strains of the El Tor

biotype (Gangarosa, 1974). One reason might be the difference between the two biotypes in the expression of virulence factors like cholera toxin. Classical strains produced about 20 times more cholera toxin in the ligated ileal loop than El Tor strains (Jonson, 1990). Classical strains can be separated from El Tor strains based on genotype, phenotype, and the severity of disease symptoms. Recently, hybrid strains have been reported (Nair et al., 2002). An example of such a hybrid is an El Tor variant that encodes the cholera toxin-encoding genes of classical strains (Safa et al., 2010).

Recent sequencing of an ancient *V. cholerae* genome from a preserved intestine of a cholera patient who died during the Philadelphia cholera epidemic of 1849 provides insights into the second *V. cholerae* pandemic. It is likely that the patient was a male who was infected as part of the second pandemic that started in 1829 and ended in 1851 (Devault et al., 2014; Pollitzer, 1959b). The resulting genome PA1849, displayed 95% similarity to the genome of the classical strain O395 it was mapped to (Devault et al., 2014). Plasmids and genomic islands unique to PA1849 and absent in O395 would not have been detected. Besides single nucleotide polymorphisms that differ between the two genomes, the PA1849 genome contained the classical *ctxB* and lacked three genomic islands (GI 11, 14, 21). Genomic islands 11 and 21 were predicted to be prophages, and GI-14 contained 12–15 genes that encode hypothetical proteins (Chun et al., 2009). The exact number of *ctx* repeats could not be determined due to low sequence coverage (Devault et al., 2014).

The T6SS of isolates from the second and sixth pandemics have not yet been systematically analyzed. To study the T6SS of a second pandemic strain, I took advantage of the available genome sequence of *V. cholerae* PA1849. I tested the competitiveness of classical strains and their ability to engage in and resist bacterial killing.

6.2 Results

6.2.1 The second pandemic strain belongs to the AAA compatibility group

Characterization of the *V. cholerae* T6SS focused on more recently isolated strains because the lack of isolates and their genome sequences made it impossible to characterize earlier strains. The reconstruction of the genome of a second pandemic *V. cholerae* strain from 1849 (Devault et al., 2014) offered a unique opportunity to study the T6SS of a second pandemic strain. Comparison between the T6SS gene clusters of a variety of *V. cholerae* strains revealed conserved genes that encode structural and regulatory components and diverse effector and immunity protein encoding genes (Unterweger et al., 2014). The latter are encoded on genetic elements we called effector modules and proposedly can be exchanged between *V. cholerae* strains. Two observations support this hypothesis. First, the GC content of these elements differs from the GC content of surrounding T6SS genes. Second, the distribution of these elements among *V. cholerae* strains does not always follow the phylogeny of these strains determined based on house-keeping genes and thus argues against inheritance of these elements by lateral gene transfer, for example for VL426. We further revealed that the set of effector modules determines the ability of a *V. cholerae* strain to coexist with other strains in direct contact under conditions that permit for T6SS-mediated killing. Two strains with the same effector module set are able to coexist and thus called compatible. Two strains that differ in their effector module set are outcompeting each other in a T6SS-dependent manner and thus called incompatible. Based on the T6SS effector and immunity proteins encoded in the gene clusters of the strain PA1849, we can make predictions about other strains PA1849 would be able to coexist with or outcompete. The integrity of the genes encoding structural and regulatory T6SS components will tell us about the functionality of the system.

To analyze the T6SS effector and immunity proteins of PA1849, we extracted the T6SS gene clusters of the reconstructed genome. Illumina reads were downloaded from the Sequence Read Archive of NCBI (accession number SRP029921). Single reads with an average length of 80 basepairs, adding up to a total of 172,138,676 raw reads, were assembled onto the T6SS gene clusters of *V. cholerae* O395 (NC_009456.1 for

chromosome 1 and NC_009457.1 for chromosome 2) using CLC software. As the DNA had been stored for 165 years, the sequences contained gaps. However, over 64% of the amino acid sequences for TsiV1, TsiV2, and TsiV3 were determined, which was sufficient to determine the compatibility group of PA1849. Alignments between the available amino acid sequences of TsiV1, TsiV2, and TsiV3 from the strain PA1849 and the El Tor strain N16961 revealed that immunity proteins of the two strains were highly similar (Figure 6-1A-C). The amino acid sequence of $TsiV1^{PA1849}$ possibly differed from that of $TsiV1^{N16961}$ at position 180 by the substitution of a tryptophan (W) with a glycine (G). To further test the confidence in this amino acid substitution, we analyzed the quality of the sequence (Figure 6-2). Based on the low coverage and the contradictory sequence of individual contigs, it was unclear whether an amino acid exchange in TsiV1 at position 180 had occurred. Over the length of the available amino acid sequence, TsiV2 and TsiV3 of PA1849 differed from the TsiV2 and TsiV3 sequences of N16961, each by one amino acid, D219E and N44D, respectively. This analysis indicated that PA1849 encodes immunity proteins of the AAA compatibility group.

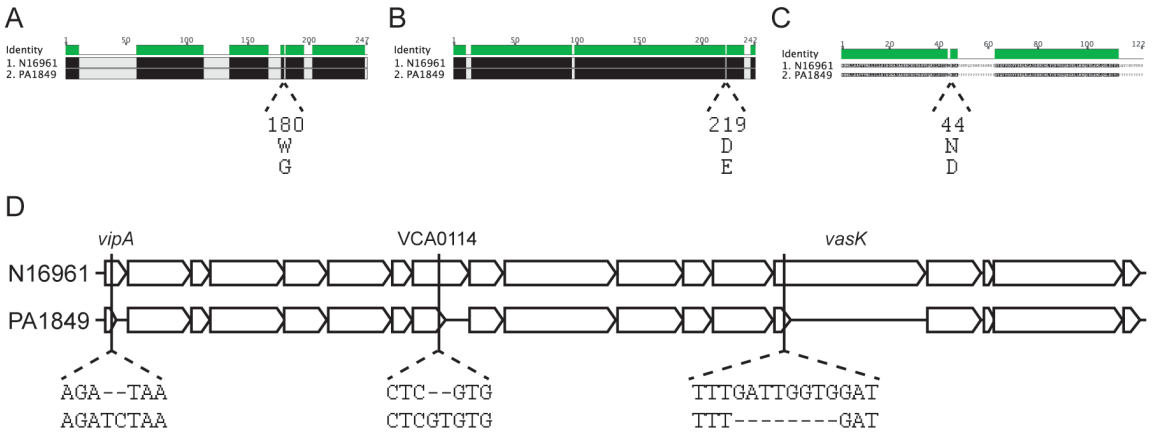


Figure 6-1 T6SS compatibility group and gene cluster of PA1849.

(A-C) The immunity proteins TsiV1 (A), TsiV2 (B), and TsiV3 (C) of PA1849 and N16961 are highly similar. MUSCLE alignments of the indicated amino acid sequences are shown. The green bar indicates identical amino acids. Three differences between the available sequences of PA1849 and N16961 lack a green bar and are highlighted and enlarged. At all the other regions for which no green bar is indicated, no sequence of PA1849 is available for comparison. (D) PA1849 encodes three nonsense mutations in the large T6SS gene cluster. Graphical depictions of the open-reading frames in the T6SS gene clusters of N16961 and PA1849 drawn to scale are shown. Nonsense mutations are highlighted and differences between the two nucleotide sequences at this point are enlarged.

Next, we asked whether PA1849 was capable of assembling a functional T6SS based on its genes encoding the proteins necessary for a functional secretion system. We compared the open reading frames in the nucleotide sequence of the large T6SS gene cluster of PA1849 to the ones in the O1 serogroup El Tor biotype strain N16961. We observed nonsense mutations in three genes (Figure 6-1D). Insertion of TC at position 174 in *vipA* resulted in a premature stop codon at position 196 instead of 507 (Figure 6-1D). Insertion of TG at position 801 in VCA0114 resulted in a premature STOP codon at position 904 instead of 1333 (Figure 6-1D). Deletion of 8 nucleotides at position 514 in *vasK* resulted in a premature STOP codon at position 528 instead of 3544 (Figure 6-1D). *VipA* encodes a component of the contractile outer T6SS sheath and a mutant lacking *vipA* does not secrete Hcp; Hcp secretion is the hallmark of a functional T6SS (Basler et al., 2012). *VasK* and *VCA0114* have been described to encode other essential components of the T6SS, as indicated by there being no detectable secretion of Hcp in their absence (Raskin et al., 2006; Zheng et al., 2011). Over the total length of the 24,160 bp that

constitute the three T6SS clusters, the nucleotide sequence could not be compared at 227 positions (0.9%) due to the lack of PA1849 sequence. In addition to the three nonsense mutations, 98 single nucleotide polymorphisms were detected, which could further affect the function of T6SS proteins if they resulted in amino acid substitutions.

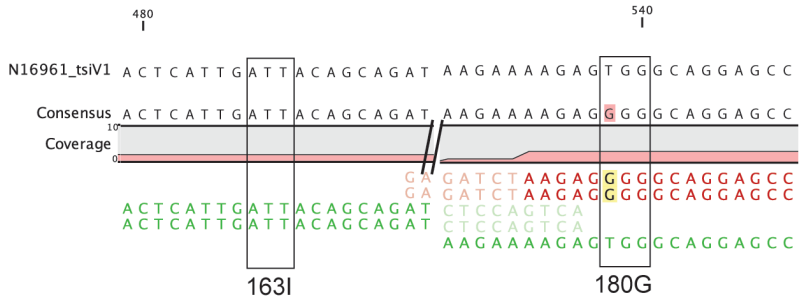


Figure 6-2 Coverage of *tsiV1* from PA1849.

Sequence reads covering the codons of the amino acids 163 and 180. The differentially coloured reads, the consensus sequence, and the sequence of N16961 are shown. The boxes highlight the codons of the indicated amino acids.

In summary, our analysis suggests that bacteria of the strain PA1849 belong to the AAA compatibility group and harbour nonsense mutations in three T6SS genes that all encode essential proteins for a functional secretion system. I concluded that PA1849 carried a nonfunctional T6SS.

6.2.2 Classical *V. cholerae* strains do not engage in microbial competition

Comparison of the whole genome of the *V. cholerae* isolate PA1849 with the genomes of isolates from more recent pandemics revealed that PA1849 is closely related to *V. cholerae* strains of the classical biotype (Devault et al., 2014). An earlier analysis of the T6SS gene cluster of the classical *V. cholerae* strain O395 also indicated nonsense mutations in the large T6SS gene cluster (MacIntyre et al., 2010). We hypothesized that nonsense mutations in the large T6SS cluster are common among classical strains spanning four cholera pandemics. We took a bioinformatics approach on a larger sample size to test this hypothesis.

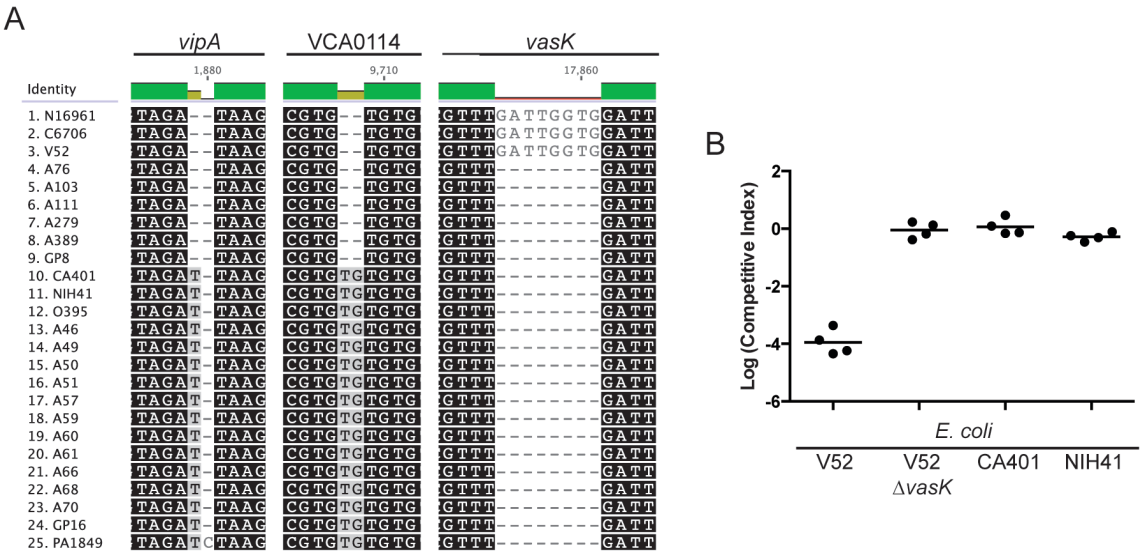


Figure 6-3 T6SS activity of classical *V. cholerae* strains.

(A) Classical *V. cholerae* strains harbor nonsense mutations in the large T6SS gene cluster. The result of a MAUVE alignment of the nucleotide sequences of the indicated strains is shown. The extractions of *vipA* start at position 170, of VCA0114 at position 800 and of *vasK* at position 510 of the individual genes. (B) Classical *V. cholerae* strains do not express Hcp when grown in liquid, on a plate, or on mucin columns. Cultures were grown under the indicated conditions and analyzed by western-blot and immunoblotting using purified antibodies against anti-DnaK and anti-Hcp serum. (C) Classical *V. cholerae* strains do not engage in T6SS-mediated killing. *E. coli* were exposed to the indicated *V. cholerae* strains in a killing assay for 4 hours at 37 ºC at a ratio of 1:1. A competitive index was determined by dividing the ratio of surviving *E. coli*/*V. cholerae* determined at t = 4 h by the ratio determined at t = 0 h. The data points and the mean of two experiments, each performed in duplicate, are shown.

To analyze the large T6SS gene clusters, we extracted the respective sequences from the genomes of 21 classical *V. cholerae* strains isolated over a 50-year period (from 1940 to 1990). The 21 sequences were aligned to the sequence of the large cluster of strain N16961 as a reference and to the sequence of the large cluster of strains C6706 and V52 that are widely used for studies of a functional T6SS (Raskin et al., 2006; Zheng et al., 2010). Sequence comparisons revealed a variable number of one to three nonsense mutations (Figure 6-3A). All 21 analyzed sequences from classical strains lacked the same eight nucleotides in *vasK* (Figure 6-3A); 16 of 22 sequences had a TG nucleotide insertion in VCA0114 and a T nucleotide insertion in *vipA* (Figure 6-3A). The same

mutations were detected in PA1849 resulting in nonsense mutations in three genes that, when deleted, abolished Hcp secretion.

To test the ability of classical strains to engage in microbial competition, we mixed the strains CA401 or NIH41 with *E. coli* in a killing assay. Streptomycin-resistant *V. cholerae* strains were mixed with rifampicin-resistant *E. coli* at a ratio of 10:1 and incubated for 4 h at 37 °C. After incubation, the bacteria were scraped off, serially diluted, and plated onto plates with streptomycin or rifampicin to select for *V. cholerae* or *E. coli*, respectively. Similar to previous experiments, killing of *E. coli* was detected when bacteria were exposed to V52 but not when exposed to V52Δ*vasK* which lacks a functional secretion system (Figure 6-3B). *E. coli* exposed to classical *V. cholerae* strains were enumerated at numbers similar to the numbers enumerated after exposure to V52Δ*vasK*. This suggested that classical *V. cholerae* strains do not engage in *E. coli* killing. Further experiments are required to test if T6SS genes are expressed under the conditions tested. The El Tor strain C6706, for example, requires an inducing signal, like the host factor mucin, to express T6SS genes under laboratory conditions.

In summary, nonsense mutations in genes encoding essential structural T6SS components were found in the genomes of all analyzed classical strains. Unlike the killing of *E. coli* observed for bacteria of the *V. cholerae* strain V52, killing of *E. coli* was not observed for the classical strains CA401 and NIH41.

6.2.3 Some classical *V. cholerae* strains are sensitive to a T6SS-mediated attack of bacteria from the same compatibility group

Compatibility groups comprise bacterial strains that have the T6SS effector and immunity proteins in common, which renders them immune to each other's T6SS-mediated attacks and allows them to coexist in direct contact (Unterweger et al., 2014). To determine the compatibility group membership of the analyzed 21 classical *V. cholerae* strains, I took a bioinformatics approach and tested compatibility experimentally.

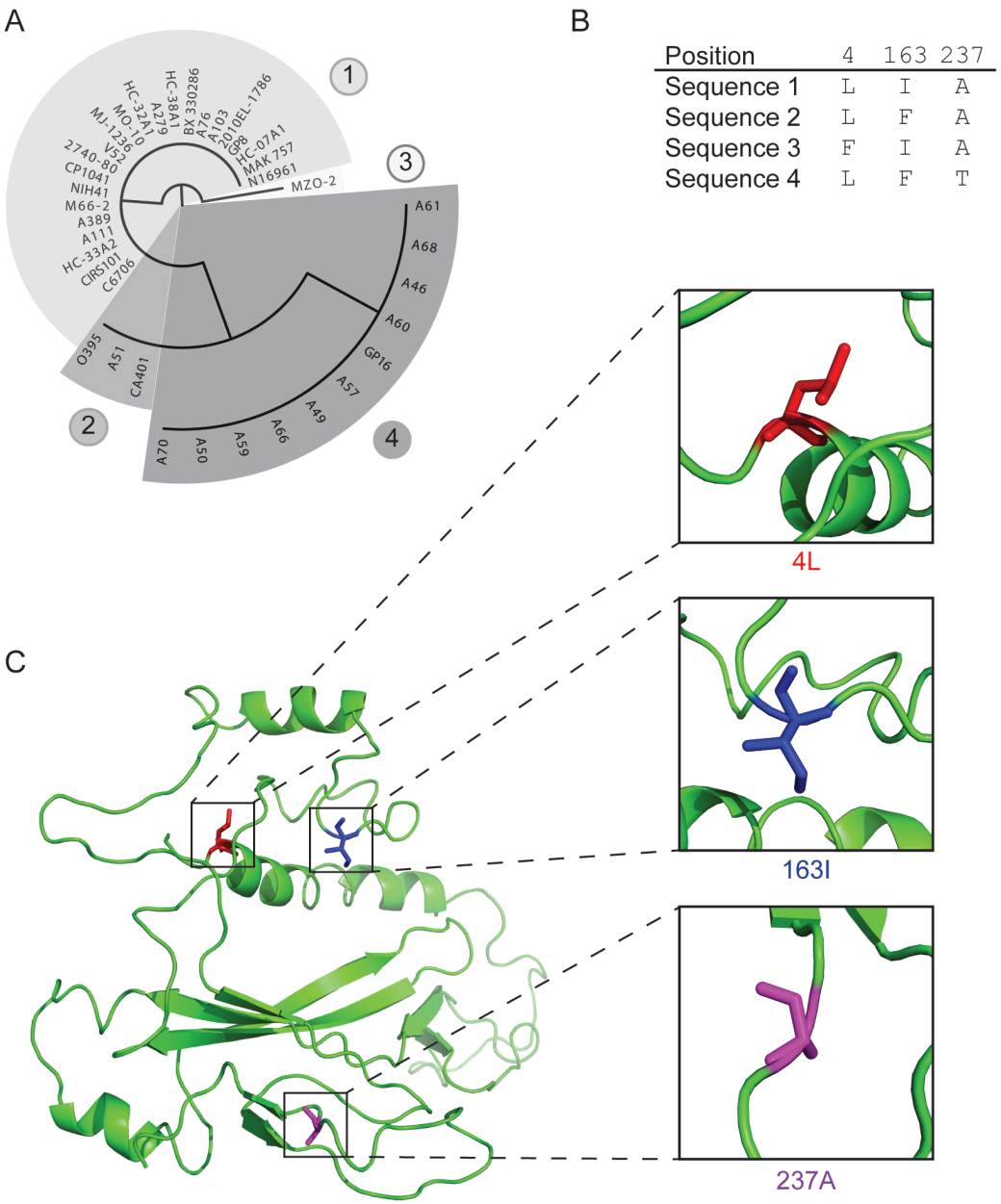


Figure 6-4 Classical strains harbor multiple amino acid substitutions in TsiV1.

(A) Amino acid sequences of TsiV1 cluster into 4 groups. A radial tree showing names of the strains that harbor the analyzed TsiV1. (B) Overview of the 4 different TsiV1 sequences. The positions and the respective amino acids at which the sequences differ from each other are shown. (C) A model of TsiV1 (sequence 1) highlighting amino acids that differ between the TsiV1 proteins. The model with up to 38% confidence was built using Phyre2 intensive modeling.

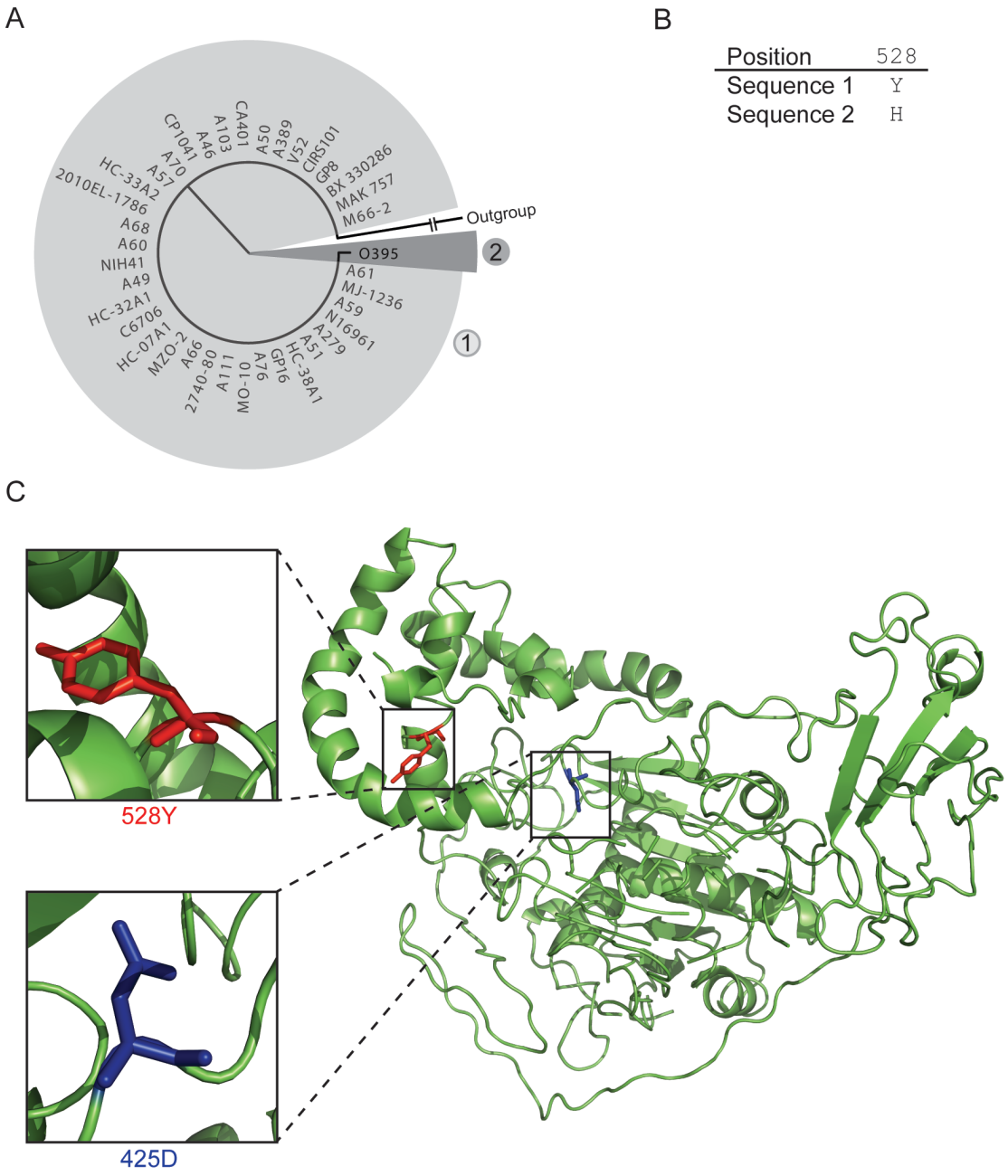


Figure 6-5 TseL is highly conserved among classical strains.

(A) TseL is highly conserved except for TseL of the strain O395. A radial tree with the names of the strains that harbor the analyzed TseL is shown. (B) TseL of O395 differs from TseL of other strains by a histidine at position 528. The position and the amino acid that differs between the TseL proteins is indicated. (C) The amino acid at position 528 maps to a different site than the active site of TseL. The model of TseL (sequence 1) was created with up to 100% confidence using intensive Phyre2 modeling.

I started my analysis with the amino acid sequences of the immunity protein TsiV1. Comparison of the TsiV1 amino acid sequences from classical strains to the TsiV1 amino acid sequences of N16961 and 16 additional strains that were previously shown to contain TsiV1 (Unterweger et al., 2014) revealed an overall sequence identity among all sequences of 98.8%. Visualization of the alignment in a radial tree shows 4 groups of sequences (Figure 6-4A). The four TsiV1 amino acid sequences vary by amino acids at positions 4, 163, and 237 (Figure 6-4B) based on single nucleotide substitutions T12G, A487T, and/or G709A in *tisV1*. To determine where the 3 amino acids map to the three-dimensional structure of TsiV1, I created a structural model of TsiV1 (sequence 1) with up to 38% confidence. The model of TsiV1 revealed that the amino acids at positions 4, 163, and 237 map to different sites of the protein. If these amino acid substitutions affect TsiV1 function, they are expected to do so differently.

TsiV1 was previously shown to inhibit TseL-mediated killing (Dong et al., 2013; Unterweger et al., 2014). To test if strains that differ in their TsiV1 amino acid sequence also contain different TseL amino acid sequences, we compared the amino acid sequences of TseL from all strains that were analyzed for TsiV1 (Figure 6-4). The alignment reveals that all strains except O395 harbor identical TseL proteins (Figure 6-5A). TseL of O395 differs by a single amino acid substitution from tyrosine to histidine at position 528 (Figure 6-5B). To determine the proximity of 528Y to the active site of TseL, which is a potential binding site for TsiV1 as shown for the interaction between TsiV3 and VgrG-3 (Zhang et al., 2014), I created a model of the protein structure of TseL with up to 100% confidence. In this model, 528Y is not found close to the active site 425D (Figure 6-5C), suggesting that this substitution does not affect the interaction with an immunity protein that usually interacts with an effector close to the active site. Taken together, the analyzed classical strains harbor TsiV1 proteins with multiple amino acid substitutions and have a highly conserved TseL.

To extend this analysis to the immunity proteins TsiV2 and TsiV3, I aligned the respective sequences of the 21 classical strains to TsiV2 and TsiV3 of N16961, respectively. TsiV2 in all the analyzed sequences is identical to TsiV2 in N16961 (Figure 6-6A). Similarly, TsiV3 in all the analyzed sequences is identical to TsiV3 in N16961 (Figure 6-6 B). Based on the similarity of TsiV1, TsiV2, and TsiV3 of the classical

strains to the respective sequences of N16961, we grouped all the classical strains into the same compatibility group AAA as N16961. Taking the different TsiV1 amino acid sequences into account that are reflected in subfamilies, the classical strains can be further differentiated into strains of the compatibility groups A1A1A1 (e.g., strain NIH41), A2A1A1 (e.g., strains CA401 and O395) and A4A1A1 (e.g., strain A46).

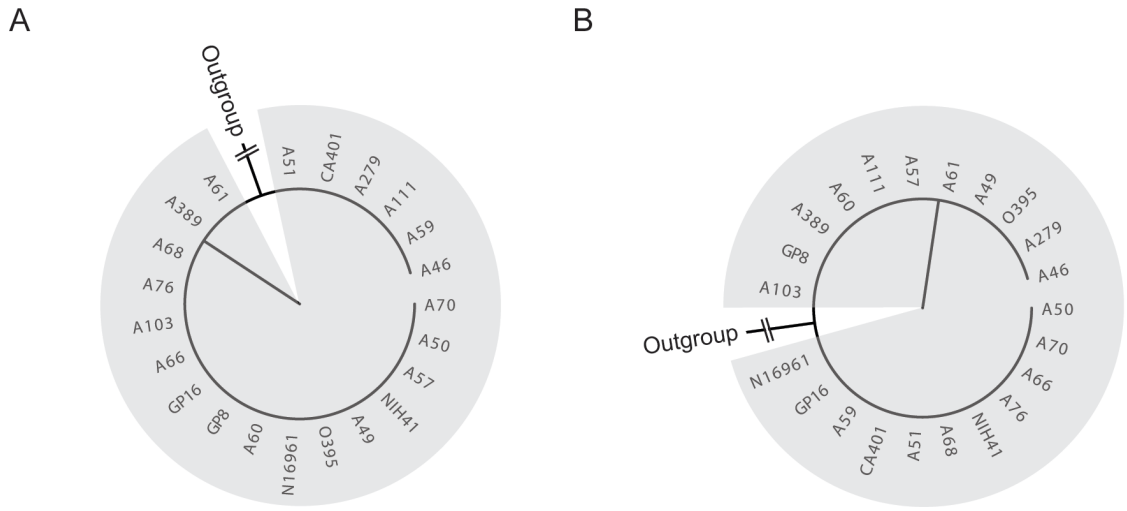


Figure 6-6 TsiV2 and TsiV3 of classical strains are highly conserved.
TsiV2 (A) and TsiV3 (B) from classical strains are identical to the respective proteins of strain N16961. Radial trees with the names of strains that harbor the analyzed sequence of TsiV2 and TsiV3, respectively, are shown. TsiV1 of N16961 was used as an outgroup.

To test experimentally if classical strains are able to coexist with other bacteria of strains in the A1A1A1 compatibility group that have an active T6SS, we exposed bacteria of the classical strains O395, CA401, and NIH41 to strain V52 and V52Δ*vasK* in a killing assay (Figure 6-7). As controls for bacteria compatible and incompatible with V52, we exposed bacteria of the strain C6706 (compatibility group A1A1A1) and AM-19226 (compatibility group C1D4F1), respectively, to V52. As expected, we recovered equal numbers of surviving C6706 when exposed to V52 or V52Δ*vasK*, whereas the number of surviving AM-19226 exposed to V52 or V52Δ*vasK* differed by 10,000,000 fold (Figure 6-7). V52 killed O395 and CA401 by 10 fold in a T6SS-dependent manner. In contrast, NIH41 was recovered in equal numbers after exposure to wild-type V52 or V52Δ*vasK*

and is thus fully immune to a T6SS-mediated attack of other bacteria of the A1A1A1 compatibility group.

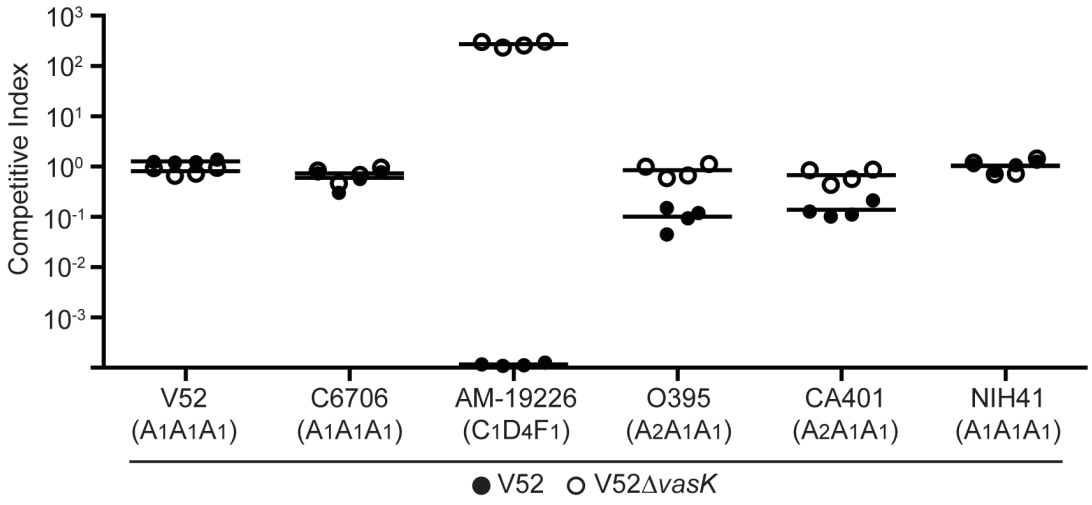


Figure 6-7 Some classical strains are susceptible to T6SS-mediated killing.

Some classical strains are sensitive to a T6SS-mediated attack of an AAA strain. V52 or V52Δ*vasK* were mixed with the indicated strains in a killing assay at a ratio of 1:1 and incubated for 4 h. The competitive index was calculated by dividing the ratio of indicated strain/V52 determined at t = 4 h by the ratio determined at t = 0 h. The mean ± standard deviation of 2 independent experiments, each performed in duplicate, is shown.

In summary, I showed that the analyzed classical strains belong to the AAA compatibility group. Two strains of the compatibility group A2A1A1 are susceptible to a T6SS-mediated attack of the A1A1A1-compatibility group strain V52.

6.2.4 A single amino acid substitution in TsiV1 causes susceptibility to a T6SS-mediated attack

To further examine the molecular basis of the sensitivity of selected classical strains to a T6SS-mediated attack of strains of the A1A1A1 compatibility group, we focused on the differences in the immunity protein TsiV1. The TsiV1 in strains O395 and CA401, which are susceptible to A1A1A1-mediated killing, differs from the TsiV1 of the fully protected strain NIH41 by the amino acid substitution of isoleucine (I) to phenylalanine (F) at position 163 in TsiV1.

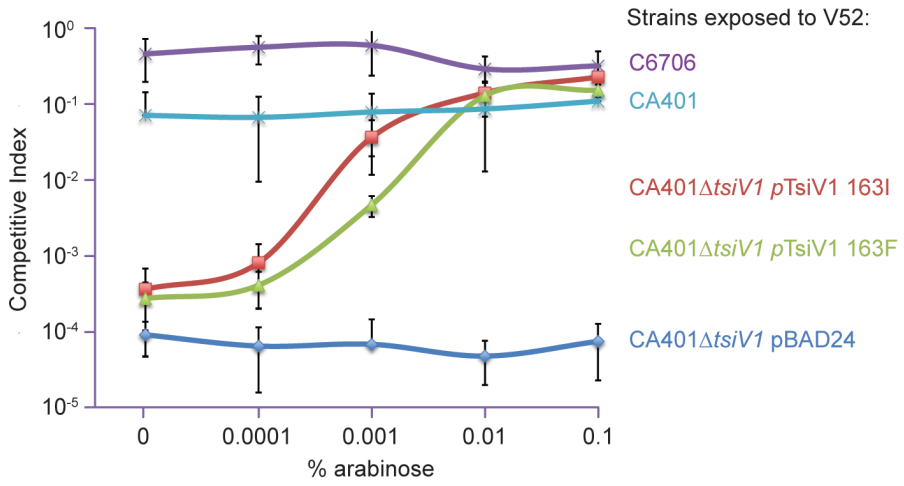


Figure 6-8 TsiV1 163I protects from a T6SS-mediated attack better than TsiV1 163F.
V52 was mixed with the indicated strains in a killing assay at a ratio of 1:1 and incubated for 4 h on LB plates supplemented with arabinose to the final indicated concentrations. The competitive index was calculated by dividing the ratio of the indicated strain/V52 determined at t = 4 h by the ratio determined at t = 0 h. The mean ± standard deviation of 2 independent experiments, each performed in duplicate, is shown.

To test the effect of this amino acid substitution at position 163 on the ability of TsiV1 to protect a *V. cholerae* strain from a T6SS-mediated attack, we performed a killing assay. An in-frame deletion mutant of *tsiV1* was created in CA401. This mutant was transformed with a pBAD24 empty vector as a control or a pBAD24 vector carrying *tsiV1* encoding either of the two alleles. CA401 and C6706 were used as controls for strains that endogenously express the proteins TsiV1 163F and 163I, respectively. A killing assay was performed in which the above strains were exposed to bacteria of the strain V52 (Figure 6-8). Similar to previous experiments, CA401 was more sensitive to T6SS-mediated killing than C6706 (Figure 6-8). A CA401 mutant lacking *tsiV1* was 3 fold more sensitive than CA401 wild-type. This indicates that the TsiV1 163F has some protective effect on CA401.

To test to what extent TsiV1 163I and TsiV1 163F rescue CA401, we induced the expression of the TsiV1 proteins exposed to V52 with increasing arabinose concentrations ranging from 0 to 0.1%. At maximal induction, both TsiV1 163I and TsiV1 163F were able to fully rescue CA401 (Figure 6-8). Under lower inducing conditions, TsiV1 163I and TsiV1 163F showed differences in their ability to rescue

CA401 (Figure 6-8). TsiV1 containing 163I was more efficient at rescuing CA401 than TsiV1 containing 163F. The low arabinose concentration of 0.001% was previously shown to induce low levels of gene expression from the pBAD promoter (Judson and Mekalanos, 2000) and mimics the naturally low expression levels of immunity protein-encoding genes as shown for the promoter upstream of *tsiV2* (Miyata et al., 2013). Sarah Miyata and colleagues created transcriptional fusions of the immunity gene promoter to *lacZ* and of the *hcp* promoter to *lacZ*. The promoter of the immunity gene drove transcription at about 8 fold lower levels than the promoter of *hcp,* as determined in a β-galactosidase assay.

Taken together, these results indicate that TsiV1 protects CA401 from a T6SS-mediated attack and suggest that when expressed at low levels, TsiV1 with isoleucine at position 163 provides better protection than TsiV1 with phenylalanine at position 163.

6.2.5 Classical strains are outcompeted by non-AAA *V. cholerae* strains

V. cholerae strains that belong to compatibility groups other than AAA are putative competitors of the classical strains. We have previously shown that strains of different compatibility groups outcompete each other (Unterweger et al., 2014). Based on the different compatibility group membership between classical strains and non-AAA strains, we hypothesized that classical strains would be outcompeted by non-AAA strains.

Therefore, we exposed classical *V. cholerae* strains to non-AAA strains DL4211, DL4215, V51, and 1587 in a killing assay at a ratio of 1:1. As a control for no competition, only streptomycin- and only rifampicin-resistant classical strains were mixed. A competitive index of 1 indicates no competition, as shown for bacteria of the classical strains exposed to kin alone (Figure 6-9). After exposure to non-AAA strains, 1,000 to 10,000 fold less bacteria of the classical strains were recovered (Figure 6-9).

Taken together, these results indicate that CA401 and NIH41 are sensitive to antiprokaryotic killing of non-AAA strains.

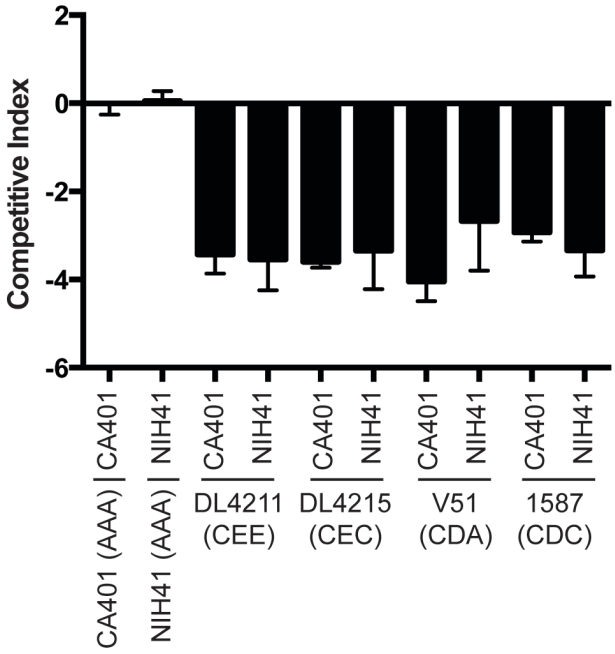


Figure 6-9 Classical strains are killed by non-AAA *V. cholerae* strains.

Classical strains CA401 and NIH41 were exposed to indicated *V. cholerae* strains at a ratio of 1:1 and incubated for 4 h at 37 °C. The competitive index was calculated by determining the ratio of CFU/ml between two strains as indicated below the graph and by dividing this ratio determined at t = 4 h by the ratio determined at t = 0 h. The mean ± standard deviation of two independent experiments, each performed in duplicate, is shown.

6.2.6 Model for gain, loss, and exchange of T6SS mutations among classical *V. cholerae* strains

Previous experiments showed genetic differences between the T6SS gene clusters of classical strains and phenotypic differences in the ability of these strains to resist and engage in microbial killing. I then became interested in how the different phenotypes relate to the various genotypes, and hypothesized that an increasing number of mutations in the T6SS gene clusters were acquired during the evolution of these strains.

Therefore, I compared the T6SS between classical and El Tor strains of a phylogenetic tree that was created by Devault and colleagues based on the core genome of the indicated strains (Devault et al., 2014) (Figure 6-10). The mutations in *vipA*, VCA0114, *vasK,* and the amino acid substitutions of TsiV1 were indicated. Strains for

which we did not previously analyze the T6SS were indicated with β. This analysis shows that the classical strains share the eight basepair deletion within *vasK* that is not found among strains of the phylocore genome-1 clade. Insertions within *vipA* and VCA0114 are found only among some classical strains. TsiV1 with a phenylalanine at position 163 are also found only among some classical strains and are not found among strains of the phylocore genome-1 clade (Figure 6-10). The analyzed changes in the genotype of the T6SS are thus unique to classical strains; the eight basepair deletion within *vasK* is the only mutation shared by all analyzed classical strains.

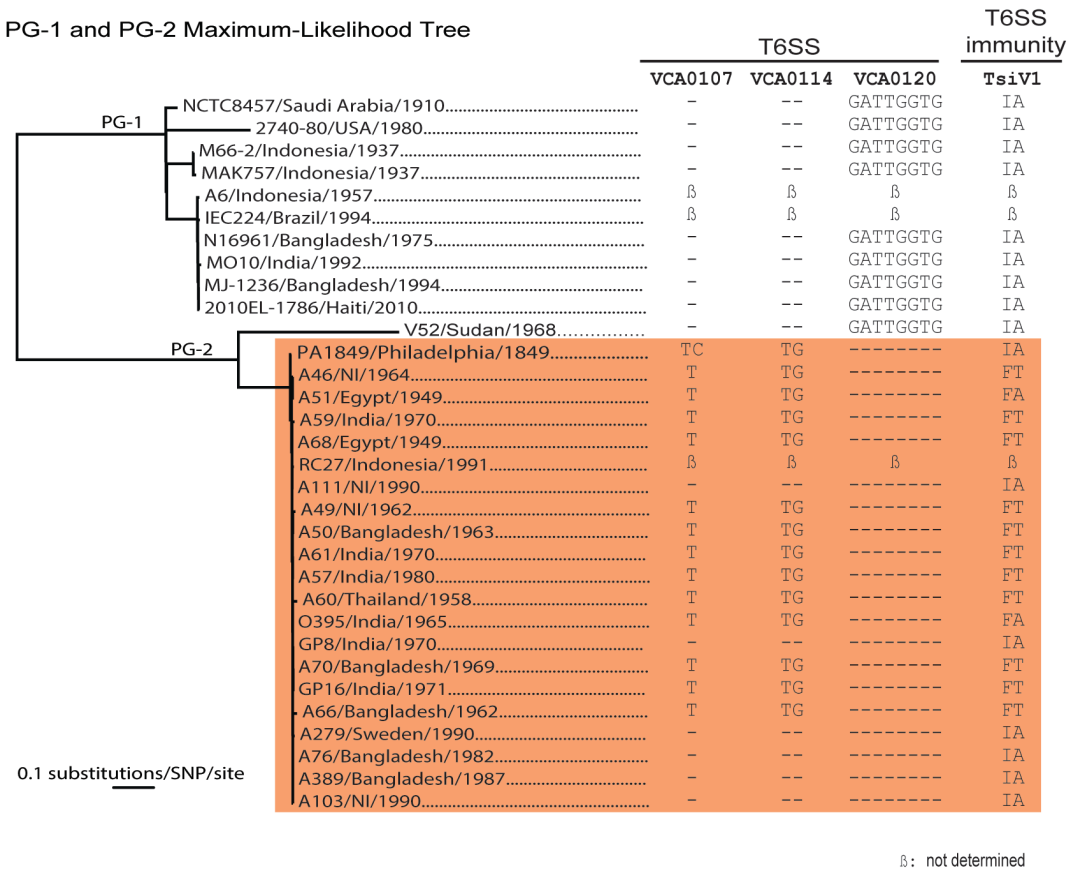


Figure 6-10 Phylogeny of *V. cholerae* strains and their T6SS gene clusters.

Maximum likelihood tree based on the core genome of indicated *V. cholerae* strains. Classical strains are highlighted in orange. Insertions, deletions, and nucleotide substitutions that were studied in previous sections are indicated on the right. Strains of which the T6SS gene clusters were not analyzed are marked with β. Phylogenetic tree reproduced with permission from Devault et al., 2014, Copyright Massachusetts Medical Society.

Figure 6-11 Phylogeny of classical *V. cholerae* strains and the acquisition of T6SS mutations.

(A) Phylogeny of classical *V. cholerae* strains. Parts of a maximum parsimony phylogeny constructed by (Devault et al., 2014) are shown. The branch length is not drawn to scale for visual clarity. Mutations within the T6SS are shown on the right. Phylogenetic tree reproduced with permission from Devault et al., 2014, Copyright Massachusetts Medical Society. (B) Model for the evolution of classical strains and the acquisition of mutations within the T6SS. Arrows indicate the described change to the T6SS gene cluster.

To study the phylogeny between closely related classical *V. cholerae* strains, we made use of a phylogenetic tree constructed by (Devault et al., 2014) in which branch length was not drawn to scale to improve visual resolution (Figure 6-11 A). Analysis of the tree and the mutations within the T6SS show that closely related strains harbor the same mutations, for example, strains A61 and A57. Examples for closely related strains are GP8, which was isolated in India in 1970 and does not contain nonsense mutations in *vipA* and VCA0114, and A70, which was isolated in Bangladesh in 1969 and contains nonsense mutations in *vipA* and VCA0114 (Figure 6-11A). Because all strains with a frameshift mutation in VCA0114, for example, share the same insertion of TG at position 801, it is unlikely that these insertions were acquired independently, and we speculate that the insertion was acquired by an ancestor these strains share in common. Identical mutations in distantly related strains are indications of horizontal gene transfer of mutated genes between distantly related strains.

To better understand the acquisition of mutations within the T6SS gene clusters during the evolution of classical *V. cholerae* strains, I built a hypothetical model based on the analyzed strains (Figure 6-11 B). In this model, the eight basepair deletion within *vasK*, the insertion of T in *vipA*, and the insertion of TG in VCA0114 were acquired by a common ancestor of all analyzed classical strains, including PA1849, and separate these strains from strains (like N16961) of the phylocore genome-1 clade. During the subsequent evolution of PA1849, separately from the other classical strains, PA1849 acquired an additional insertion of cytosine in *vipA* (Figure 6-11B). The common ancestor in strains other than PA1849 acquired a nucleotide substitution A487T in *tsiV1,* causing the exchange of isoleucine with phenylalanine at position 163 of TsiV1 and giving rise to strains like A51. Strains like A66 would have resulted from an additional nucleotide substitution of G709A in *tsiV1*. Strains that are the furthest evolved (like A76) do not show signs of nonsense insertions in *vipA* and VCA0114, or signs of nucleotide substitutions in *tsiV1*. This could be explained by horizontal gene transfer events between these classical strains and strains with an intact T6SS gene cluster like N16961 (Figure 6-11B). In summary, our model suggests that classical strains acquired mutations in their T6SS gene clusters in multiple steps. Mutations in genes encoding structural T6SS genes are suggested to predate mutations in the immunity protein encoding genes. The

individual mutations possibly were inherited from a common ancestor in most cases and might have been reversed in selected cases by the acquisition of intact genes via horizontal gene transfer.

6.3 Discussion

Starting with the analysis of the T6SS gene cluster of the oldest sequenced classical *V. cholerae* isolate, we used bioinformatics, genetics, and functional assays to study the competitiveness of classical O1 serogroup *V. cholerae* strains. Our results suggest that the second pandemic strain PA1849 belongs to the AAA compatibility group and carries nonsense mutations in its large T6SS gene cluster. The same mutations that affect genes encoding essential T6SS components are found in the genomes of 21 additional classical strains. Analysis of classical strains in a killing assay showed that they do not engage in T6SS-mediated competition. Even though all analyzed classical strains belong to the AAA compatibility group, a subset of strains carries an amino acid substitution in TsiV1 rendering these strains susceptible to a T6SS-mediated attack of an AAA compatibility group strain with an active T6SS. Like all AAA compatibility group strains tested so far, our results indicate that classical strains are susceptible to antibacterial killing by non-AAA strains. Finally, we wondered how mutations in T6SS genes were acquired during the evolution of *V. cholerae* strains and speculated about the acquisition and the exchange of mutated T6SS genes among classical strains.

Based on our results, we propose a model for the competitiveness of classical O1 serogroup strains in T6SS-mediated microbial competition (Figure 6-12). In the remaining discussion, I focus on three key aspects of microbial competitiveness of classical strains: the ability to engage in T6SS-mediated killing, susceptibility to killing by other *V. cholerae* strains, and coexistence with strains of the same compatibility group.

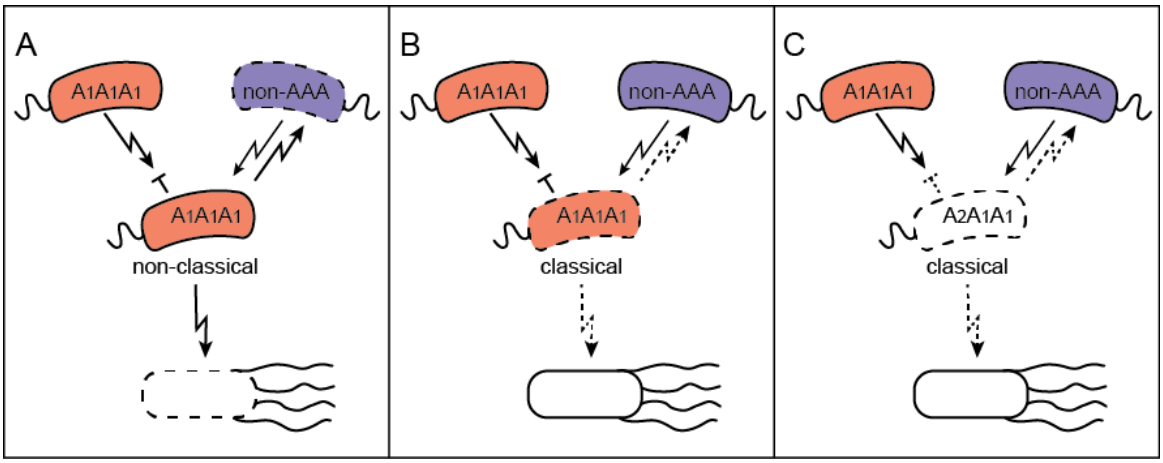


Figure 6-12 Model summarizing the competitive behaviour of non-classical and classical *V. cholerae* O1 serogroup strains.

(A) Nonclassical strains of the A1A1A1 compatibility group contain a functional T6SS to kill non-AAA *V. cholerae* strains (top right) and bacteria of other species (bottom). Fully functional immunity proteins protect from a T6SS-mediated attack of other bacteria of the A1A1A1 compatibility group (top left). (B) Inability to engage in T6SS-mediated killing renders strains like NIH41 sensitive to a T6SS-mediated attack of non-AAA strains. (C) Mutations in immunity proteins like TsiV1 renders strains like CA401 additionally sensitive to a T6SS-mediated attack of strains of the A1A1A1 compatibility group.
Arrows indicate the direction of microbial killing. The same colour indicates the same compatibility group. Dotted lines indicate the inability to engage in killing or the susceptibility of bacteria to killing by other bacteria.

Our results suggest that classical strains do not utilize their T6SS to kill other prokaryotic cells under laboratory conditions (Figure 6-12). This lack of killing could originate from the lack of expression of T6SS genes or the inability of expressed T6SS proteins to form a functional T6SS apparatus. Further experiments including qRT-PCR analysis of T6SS genes would be required to determine the expression levels of T6SS genes in classical strains. If classical strains do not express T6SS genes under laboratory conditions, we could pursue experiments looking for Hcp expression in the presence of known activators of the T6SS, like mucin (V. Bachmann and B. Kostiuk, unpublished observation), expression of VasH (Kitaoka et al., 2011), and expression of Tfox (Borgeaud et al., 2015). To test the cause-effect relationship between nonsense mutations in the three genes *vipA*, VCA0114, *vasK*, and the inability to form a functional T6SS, I

suggest the repair of these mutations on the chromosomes of classical strains and the testing of their ability to gain a functional T6SS by analysing Hcp secretion and killing of *E. coli*. Alternatively, the three nonsense mutations could be inserted into the chromosome of V52, which has an active T6SS. The resulting mutants could be analyzed for their ability to secrete Hcp as read-out for an active T6SS. Loss of Hcp secretion upon insertion of the nonsense mutations would indicate that the mutation interferes with a protein required for a functional T6SS.

The inability to engage in bacterial killing has advantages and disadvantages. Advantages are an increase in available energy resources and protection from an immune response. The T6SS requires ATP for its disassembly (Pietrosiuk et al., 2011) that could be used for other purposes if a strain is not having an active T6SS. Experiments on the infant mouse model of cholera have shown that the T6SS activates an immune response (Ma and Mekalanos, 2010). The absence of T6SS activity could allow a bacterium to evade the immune system during infection. In contrast, the inability to kill microbial competitors might lower the availability of foreign DNA for uptake (Borgeaud et al., 2015), reduce resistance to conjugation (Ho et al., 2013) and lower the response to antiprokaryotic activities of other bacteria (LeRoux et al., 2015). In summary, we observe that an inability to engage in microbial killing reduces the competitiveness of classical strains (Figure 6-12 A, B interaction with bacterium on the bottom).

During intraspecific competition in which bacteria of two strains from different compatibility groups fight against each other in a T6SS-dependent manner, the ability to engage in T6SS-dependent competition can be crucial to win the fight (Unterweger et al., 2014). For example, when incubating bacteria of the AAA compatibility group strain V52, that have an active T6SS, with bacteria of the non-AAA compatibility group strain AM-19226 at a ratio 1:1, after four hours 1,000 fold more V52 than AM-19226 is recovered (Unterweger et al., 2014). When incubating V52Δ*vasK*, that does not have an active T6SS, with AM-19226 at a ratio 1:1, after four hours 1,000 fold less V52Δ*vasK* than AM-19226 is recovered. Our experiments showed that the classical strains are outcompeted by non-AAA *V. cholerae* strains with an active T6SS. To test if an active and functional T6SS would rescue the classical strains in such a scenario, we could make use of the classical strains with a repaired T6SS. In our model, the lack of classical *V.*

cholerae strains to counterattack T6SS-mediated killing by other *V. cholerae* strains further restricts their competitiveness (Figure 6-12 A, B, interaction with bacterium top right).

The majority of the classical strains analyzed here encode an amino acid substitution in TsiV1 that increases their sensitivity to a T6SS-mediated attack of *V. cholerae* strains from the AAA compatibility group. Coexistence with other bacteria of the same compatibility group is an important feature of the concept of compatibility. The difference in this phenotype subdivides the analyzed classical strains into two groups (Figure 6-12 B, C). The inability of most classical strains to coexist with A1A1A1 compatibility group strains with an active T6SS further narrows the spectrum of bacteria with which classical strains with the TsiV1 163F protein can coexist in direct contact. Pradhan and colleagues observed that El Tor strains outgrow classical strains in coculture but not in monoculture (Pradhan et al., 2010). When grown in direct contact with the El Tor strain N16961 for over 6 h in liquid culture, the classical strain O395 enters a culturable but nonviable state (Pradhan et al., 2013). These observations indicate that classical strains are outcompeted by strains of the A1A1A1 compatibility group also under conditions in which T6SS-mediated killing does not occur, such as liquid culture (MacIntyre et al., 2010).

I summarize three forms of competitiveness among *V. cholerae* strains based on the ability to engage in and respond to a T6SS-mediated attack. Classical strains lack the ability to engage in a T6SS-mediated attack (Figure 6-12 B, C) and lack the ability that other strains have to kill bacteria of another or the same species (Figure 6-12 A). Most of the analyzed classical strains additionally lack protection from a T6SS-mediated attack of bacteria of the A1A1A1 compatibility group (Figure 6-12 C). Our model suggests that mutations affecting the ability to engage in T6SS-mediated killing predate the acquisition of mutations affecting the ability to be protected from a T6SS-mediated attack of the A1A1A1 compatibility group.

If classical strains are such weak microbial competitors, how then can they successfully have caused six cholera pandemics? One possible explanation is a scenario in which classical strains outcompete their microbial competitors by means other than killing them. Classical strains are successful in infecting the human host and in causing

severe disease symptoms therein (Gangarosa, 1974). For strains of the El Tor biotype, which produce less cholera toxin and cause milder disease symptoms than classical strains, the elimination of bacterial competitors by killing them in a T6SS-dependent manner might be important for the following reason. A certain level of cholera toxin is required to induce water efflux from intestinal epithelial cells. Because El Tor strains produce cholera toxin at low levels, high numbers of bacteria might be required to locally produce enough cholera toxin to reach the water efflux threshold. T6SS-mediated killing might be one mechanism for El Tor strains to locally reach a high cell density of cholera-toxin producing bacteria.

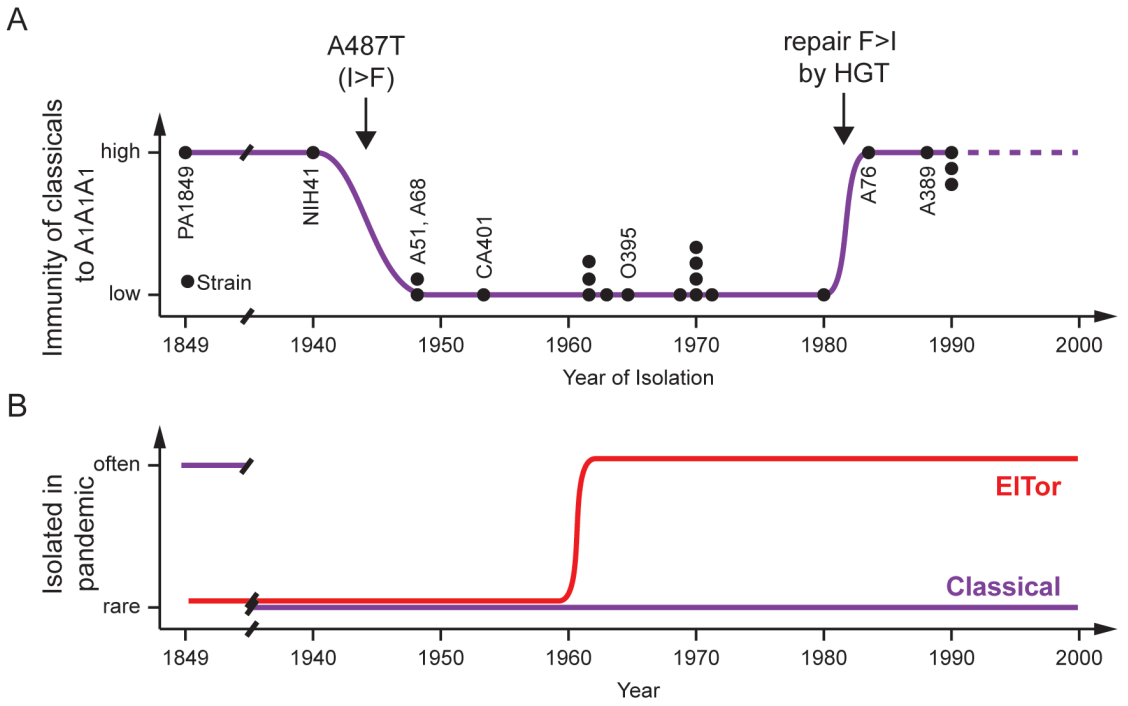


Figure 6-13 Classical *V. cholerae* strains and cholera pandemics.

(A) Classical *V. cholerae* strains (indicated as dots) ordered by year of isolation (x-axis) and the amino acid at position 163 of TsiV1 (y-axis). The nucleotide change from adenosine to thymine at position 487 of *tsiV1* that gives rise to TsiV1 163F with an impaired ability to protect against TseL of A1A1A1 compatibility group strains is indicated by an arrow. The second arrow indicates the timepoint after which classical strains with TsiV1 163I, possibly as a result of a horizontal gene transfer (HGT) event, were isolated. (B) The 7th cholera pandemic starting in 1961 is caused by O1 serogroup strains of the El Tor biotype. Strains of the classical biotype are only isolated occasionally.

Our model for the acquisition of mutations in the T6SS gene cluster indicates a horizontal gene transfer event that restores *vipA*, VCA0114, and *tsiV1*. Despite the acquisition of intact *vipA* and VCA0114, the eight basepair deletion in *vasK* and thus a dysfunctional T6SS remain. In contrast, the acquisition of TsiV1 with an isoleucine at position 163 affects the phenotype of the strain by restoring its ability to coexist with strains of the A1A1A1 compatibility group. Six classical strains that were isolated between 1982 and 1990 benefitted from this horizontal gene transfer event (Figure 6-13A). Fifteen classical isolates harboring a truncated TsiV1 with a phenylalanine at position 163 were isolated in the years 1949 to 1980. The strains isolated prior to 1949 contained isoleucine at position 163 and a fully protective TsiV1. This finding further supports the acquisition of an amino acid substitution in TsiV1 in bacteria that already contained a dysfunctional T6SS. Because the cognate effector of TsiV1, TseL, is highly conserved (Figure 6-5), amino acid substitutions in TsiV1 would have not been tolerated in bacteria with a functional T6SS, and bacteria harboring such mutations would have been killed immediately by their parent or sister cells. The ability to coexist with bacteria of the A1A1A1 compatibility group might have become more important for classical strains over time because El Tor strains of the A1A1A1 compatibility group caused the 7th pandemic starting in 1961 (Figure 6-13B). El Tor strains thus arose as a pandemic during a time in which the isolated classical strains were very weak microbial competitors and susceptible to a T6SS-mediated attack of A1A1A1 compatibility group strains like those of the El Tor biotype.

Even though the second pandemic strain PA1849 was likely unable to engage in T6SS-mediated killing, our analysis suggests that it belonged to the AAA compatibility group. The concept of compatibility groups is based on effector modules that we hypothesize can be exchanged between *V. cholerae* strains by horizontal gene transfer. The presence of the AAA effector modules in the second pandemic strain indicates that these modules were acquired by *V. cholerae* prior to 1849. Based on the presence of the AAA effector modules in pandemic strains spanning the second to the current 7th pandemic, it is tempting to speculate that the acquisition of any other effector modules did not result in a selective advantage for pandemic strains. This could be due to a

temporarily dysfunctional T6SS and due to a selection pressure common to all pandemic strains to retain the AAA effector modules.

In summary, our results suggest that classical *V. cholerae* strains that lack the ability to attack and to defend themselves are weak microbial competitors compared to nonclassical strains. The difference in microbial competitiveness could be one of numerous factors that potentially played into the replacement of pandemics caused by classical strains with the 7th and currently ongoing cholera pandemic caused by strains of the El Tor biotype.

Chapter 7
General Discussion

7 General Discussion

7.1 Diversity of T6SS effectors of *V. cholerae*

Our analyses revealed a diversity of T6SS effectors that enable *V. cholerae* to outcompete prokaryotic and eukaryotic competitors. A variety of effectors can be found within and between *V. cholerae* strains. Here I discuss the biological significance of effector diversity and the T6SS, system requirements that are necessary to translocate such a diversity of effectors.

7.1.1 Role of T6SS effector diversity in T6SS-mediated killing

T6SS effectors differ in their enzymatic activity and substrate specificity. The presence of the respective substrate in the target cell and conditions that favor a detrimental effect of the effector enzymatic activity on the target cell are required for an effector to kill a bacterium.

As of today, five T6SS effectors with four different enzymatic activities have been characterized in the *V. cholerae* strain V52. This O37 serogroup strain was used to identify T6SS effectors because its T6SS gene cluster is similar to that of pandemic *V. cholerae* strains and it has an active T6SS under laboratory conditions. The five characterized effectors comprise TseL with lipase activity, VasX with pore-forming activity, VgrG-1 with actin-crosslinking activity, and VgrG-3 and TseH with peptidoglycan-degrading activity (Altindis et al., 2015; Brooks et al., 2013; Miyata et al., 2013; Pukatzki et al., 2007; Russell et al., 2013). VasX and TseL are referred to as cargo effectors because they are not part of VgrG proteins. The substrate specificity of VgrG-3 and TseH needs to be further examined, as a wide variety of peptidoglycan-degrading effectors with varying substrate specificities have been described (Russell et al., 2011). *V. cholerae* strains other than V52 contain a variety of effectors with similar putative enzymatic activities (Unterweger et al., 2014). Effectors that are predicted to have the same enzymatic activity but differ in amino acid sequence by more than 70 percent are likely to have different substrate specificities (Unterweger et al., 2014). Such differences in substrate specificity have been described for peptidoglycan-degrading and lipase effectors (Russell et al., 2011; Russell et al., 2013). A T6SS effector with DNAse activity

has been identified in the *V. cholerae* strain HE48 (Ma et al., 2014b). The variety of enzymatic activities and substrate specificities among T6SS effectors reflects the variety of cell types and cellular structures found in nature.

Individual T6SS effectors target a defined spectrum of cells. For example, the actin-crosslinking domain of VgrG-1 cause actin-crosslinking in eukaryotic cells but does not harm prokaryotic cells (MacIntyre et al., 2010; Pukatzki et al., 2007). No cognate immunity protein to the ACD of VgrG-1 has yet been described. Because a cognate immunity protein has been identified for all the antiprokaryotic effectors in *V. cholerae*, it is unlikely that the ACD of VgrG-1 plays a role in prokaryotic killing. This notion is further supported by the results of a killing assay in which equal levels of T6SS-mediated killing of an ACD-deficient mutant of V52 and wild-type V52 were observed (MacIntyre et al., 2010). The ability of *V. cholerae* V52 to kill a variety of *V. cholerae* strains and other bacterial species like *E. coli* has been tested by exposing the strain of interest to wild-type V52 and V52Δ*vasK* with a dysfunctional T6SS (MacIntyre et al., 2010; Unterweger et al., 2014). The role of individual effectors in the killing of individual bacterial species still remains to be studied. Killing assays in which V52 was mixed with *E. coli* or C6706Δ*tsiV1-3* provide first indications for differing roles of individual effectors in the dependence of an attacked bacterium. For example, VgrG-3-mediated killing had a stronger effect on killing of *E. coli* than *V. cholerae* C6706 whereas VasX-mediated killing of *E. coli* was less than that of *V. cholerae* C6706 (chapter 2). Despite the ability of *V. cholerae* V52 to kill a wide range of bacterial species, its target range is limited. No killing of *P. aeruginosa* has been observed (MacIntyre et al., 2010) despite the lack in *P. aeruginosa* of immunity proteins homologous to those in *V. cholerae*. Also, *V. cholerae* has not yet been shown to kill any Gram-positive bacteria (MacIntyre et al., 2010). One reason for the inability of *V. cholerae* V52 to kill Gram-positive bacteria, could be that peptidoglycan is located outside of Gram-positive bacteria. T6SS effectors are likely not getting access to this peptidoglycan because they are injected into the cytoplasm of Gram-positive target cells.

In addition to a limited target range, T6SS effectors require certain conditions to be deadly. Fu and colleagues observed killing of infant rabbits by VgrG-3 but not by VasX (Fu et al., 2013). Sarah Miyata observed VasX-mediated killing under conditions

of low but not high osmolarity (Miyata et al., 2013); likely due to the pore-forming activity of the effector and subsequent water influx into the cell in an environment of low osmolarity. Such an environment of low osmolarity might not be present in the intestine. Competition experiments between *P. aeruginosa* and *Agrobacterium tumefaciens* further exemplify the role of the surrounding conditions in the outcome of bacterial killing (Ma et al., 2014a). Ma and colleagues coincubated the two species on a LB agar plate and observed that *P. aeruginosa* outcompetes *A. tumefaciens*. When repeating the same experiment on a plant leaf, *P. aeruginosa* was outcompeted by *A. tumefaciens*. The role of individual effectors might explain the environmental influence on the outcome of this experiment, and environmental conditions can possibly differentially regulate the expression of individual effectors within a single bacterium. Further experiments are needed to test this hypothesis.

These observations indicate that the diversity of T6SS effectors reflects the variety of molecular structures in competing cells. The composition of T6SS effectors therefore determines the target range of a given *V. cholerae* strain and the conditions under which it will be competitive.

7.1.2 System requirements of the T6SS to accommodate the diversity of effectors

The T6SS requires certain features to translocate such a diversity of effectors. These features are required for a single secretion system of a given strain to translocate multiple diverse effectors and for strains to incorporate newly acquired effectors into their T6SS. I assume that the same set of effectors is translocated with each firing event of a bacterium. It is also possible that the translocated effector set differs between firing events, either because individual effectors are differentially regulated or because certain effectors are present in excess and occupy the binding sites of effectors present at low abundance.

The T6SS of *V. cholerae* provides three conserved interfaces to load up to five different effectors at once onto VgrG proteins at the tip of the T6SS (Figure 7-1). One interface is a covalent bond between the conserved core of a VgrG protein and an enzymatic effector domain as a C-terminal extension. Examples in strain V52 are the

actin-crosslinking domain of VgrG-1 and the peptidoglycan-degrading domain of VgrG-3 (Brooks et al., 2013; Pukatzki et al., 2007).

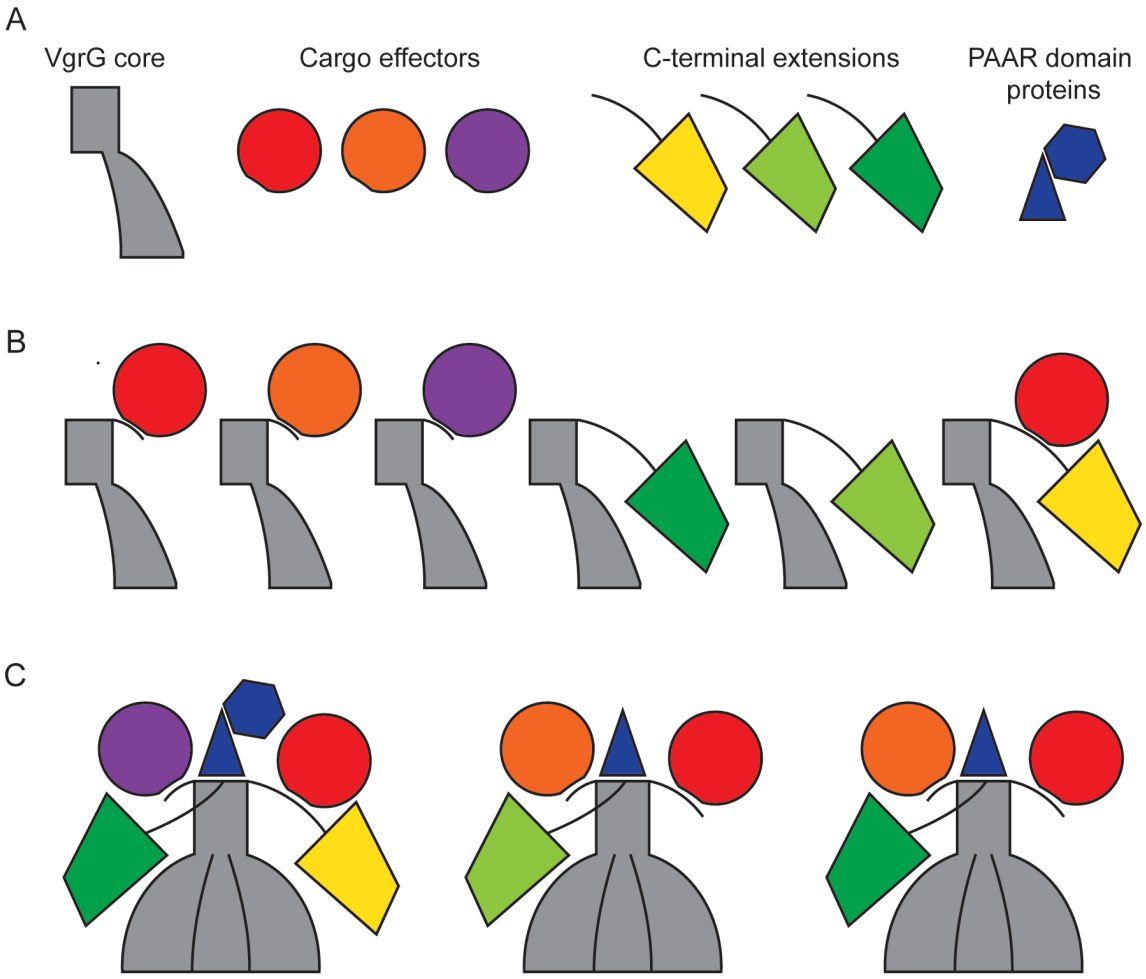


Figure 7-1 Simplified model of the loading of diverse T6SS effectors.

(A) *V. cholerae* strains encode a conserved VgrG core and a variety of cargo effectors, C-terminal extensions of VgrG proteins, and PAAR-domain proteins. (B) Cargo effectors and C-terminal extensions can be combined with the VgrG core. The interaction sites between cargo effectors and VgrG proteins remain to be identified and are speculatively indicated in this cartoon. (C) Three VgrG proteins form the trimer at the tip of the T6SS. Three examples for T6SS systems loaded with a different number and different kinds of effectors are shown. A variety of effector proteins can be associated with this tip, depending on what proteins are harboured by the individual strain.

Another interface is provided for cargo effectors like VasX and TseL by adaptor proteins like VasW and Tap-1, respectively, that likely load these effectors onto the secretion system (Miyata et al., 2013; Unterweger et al., accepted). The putative interaction site(s) at which VgrG proteins interact with cargo effectors remains to be identified. A third interface is a conserved domain of VgrG that mediates the interaction with PAAR domain proteins located at the very tip of the secretion system (Shneider et al., 2013).

Similar enzymatic activities are found in effectors that use different interfaces. For example, peptidoglycan-degrading effectors can be cargo effectors or C-terminal extensions of VgrG proteins (Unterweger et al., 2014). However, the large size of an effector and its subsequent steric hindrance to other T6SS components could be a reason for the large pore-forming effector VasX (121 kDa), for example, to be rather a cargo effector than a C-terminal extension of a VgrG (size of the ACD of VgrG-1 is 53 kDa). Adjustments of the amino acid sequence in the form of positive and negative selection (chapter 5) might be necessary to fine-tune possible interferences between effectors to create a stable effector complex at the tip of the secretion system. I assume that effectors attach to the tip of the T6SS independently of each other, because no direct interaction between individual cargo effectors or cargo effectors and C-terminal extensions of VgrG proteins with enzymatic activities or PAAR domain proteins has been reported.

The common use of these interfaces enables *V. cholerae* strains to incorporate a variety of effectors onto the same secretion apparatus. Theoretically, one secretion system could have as many as three VgrG-proteins with 3 C-terminal extensions, three cargo effectors and at least one PAAR domain effector. It remains to be elucidated if the T6SS of *V. cholerae* could evolve to carry that many effectors or if there is a limit to the number of effectors that can be translocated with one firing event.

7.2 Genetic organization of T6SS gene clusters in *V. cholerae*

Most of the T6SS proteins are encoded in genes on one of three T6SS gene clusters on the small and large *V. cholerae* chromosome. We described variable effector modules among otherwise conserved genes of the clusters (Unterweger et al., 2014). In

this chapter, I discuss the evolutionary origin of the T6SS gene cluster as a whole, the current diversity of effector modules, propose a model for the evolution of the auxiliary gene cluster 1, and discuss the genetic mechanisms putatively required to assemble such a T6SS gene cluster.

7.2.1 A possible phage origin of the T6SS gene cluster

V. cholerae strains analyzed so far all encode large and auxiliary clusters on the small and large chromosomes. Our alignments showed that genes encoding structural components are highly conserved (Unterweger et al., 2014). Structural proteins encoded by T6SS genes have homology to proteins of tailed bacteriophages, as indicated by the following examples. A bioinformatics approach revealed significant sequence homology between VgrG and gp27 and gp5 proteins of phage T4 (Pukatzki et al., 2007). Superimposition of the crystal structure of Hcp onto the crystal structure of gp27 tube proteins shows that the two proteins are very similar (Leiman et al., 2009). Shneider and colleagues identified genes in the T6SS gene clusters that encode proteins with a PAAR domain similar to the T4 bacteriophage protein gp5.4 (Shneider et al., 2013). Basler and colleagues compared the outer T6SS sheath proteins VipA and VipB with the T4 phage sheath protein gp18 and found conserved domains in phage sheath and T6SS sheath proteins (Kudryashev et al., 2015). These findings suggest a common evolutionary origin of the T6SS and phage components. In support of this idea, the GC content of the T6SS gene cluster in *V. cholerae* differs from the overall GC content of the whole genome suggesting that *V. cholerae* acquired the T6SS at some point in the past. It is also possible that the T6SS evolved by convergent evolution.

Although multiple phage proteins and T6SS components share similarities, many of their proteins are different. An example of a protein unique to the T6SS is the Icm-family protein TssM in the membrane anchoring complex that shares homology with a similar protein of type IV secretion systems (Bingle et al., 2008; Cascales, 2008). Unique to bacteriophages are the phage head or tail fibres (Veesler and Cambillau, 2011). Also, phages use integrases to insert their genome into the chromosome (Groth and Calos,

2004). Integrases have not yet been associated with T6SS genes that encode structural proteins of the secretion system apparatus.

7.2.2 The T6SS harbors mobile effector modules

Whereas the genes encoding structural and regulatory proteins of the T6SS are highly conserved among *V. cholerae* strains (Unterweger et al., 2014), the T6SS gene clusters of the *V. cholerae* strain N16961 contain multiple genes that differ or are missing in other strains (Figure 7-2; Appendix A Figure 9-2). The presence and absence of these genes in various T6SS gene clusters and the difference in their GC content compared to the surrounding genes of the T6SS gene cluster suggests that these genes are encoded on mobile elements. We call such mobile elements "effector modules" (Unterweger et al., 2014) and they differ between *V. cholerae* strains, as explained below.

The auxiliary cluster 1 differs between *V. cholerae* strains by the presence or absence of an actin crosslinking domain (ACD) in *vgrG-1* and by effector modules that span three genes. The genes of this effector module encode the C-terminus of the adaptor protein Tap-1, an effector protein with antiprokaryotic activity, and an immunity protein (Figure 7-2). The auxiliary cluster 2 differs among strains by effector modules spanning four genes. Genes of this module encode a short C-terminal extension of VgrG-2, an adaptor protein, an effector protein, and an immunity protein (Figure 7-2). The large cluster differs by unique elements at its 3' end which span two genes. These genes encode the C-terminal extension of VgrG-3 with effector activity and a cognate immunity protein. The large cluster also encodes 2 genes (VCA0105, VCA0106) at its very 5' end; one of these genes encodes a PAAR domain protein and might be regulated separately from the remaining genes of the large cluster. Bernard and colleagues identified a putative σ^{54} binding sequence downstream of VCA0106 and upstream of *vipA* (Bernard et al., 2011). Even though all analyzed strains contain the two genes VCA0105 and VCA0106 (Appendix A, Figure 9-3), they might be part of an element that is mobile and thus might be missing from some strains. Another PAAR domain protein is encoded in a six gene-long element outside the three T6SS gene clusters on the small chromosome. Other genes on this element encode the T6SS effector TseH, its immunity protein TsiH, a

putative integrase, an IS5 transposase, and a hypothetical protein (Appendix A, Figure 9-4). This element is present in all analyzed pandemic strains but is found only occasionally in nonpandemic strains (Appendix A, Figure 9-2, 9-4, 9-5). Taken together, effector modules all encode at least the part of an effector protein with enzymatic activity and an immunity protein but differ in length and in the total number of genes they contain.

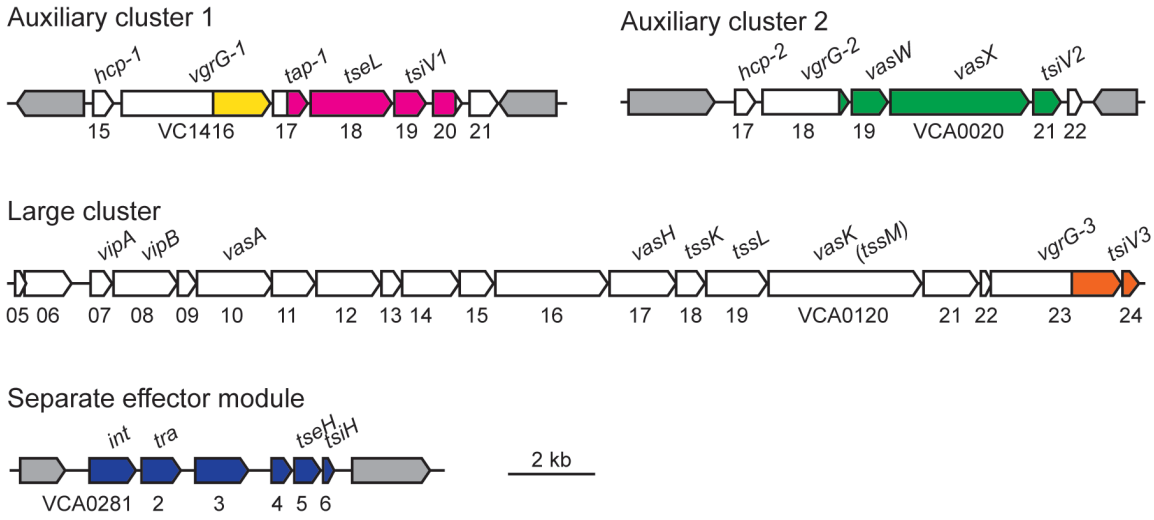


Figure 7-2 Overview of T6SS effector modules in *V. cholerae* T6SS gene clusters.
Graphical depiction of the T6SS gene cluster of the *V. cholerae* strain N16961. All genes or part of genes that are exchanged by other genes or completely missing in other strains are shown in colour. The gene labeled "int" encodes a putative integrase, the gene labeled "tra" encodes a putative transposase.

TseH is an effector that contains a characteristic YDxxGRL(I/D) repeat motif that is common to a variety of Rhs proteins (Altindis et al., 2015; Hill, 1999). The genes encoding Rhs proteins were described in *E. coli* before the T6SS effector function of the proteins was known. Some *rhs* genes were found directly downstream of *vgrG* genes, unlike *tseH* in *V. cholerae,* which is not part of the three T6SS gene clusters but encoded in a separate effector module on the small chromosome (Figure 7-2). Repetitive gene duplication events were observed next to *rhs* genes in *E. coli* and flanking regions were subsequently identified as "rearrangement hot spots" (Lin et al., 1984). Experiments further examining recombination events next to *rhs* genes revealed that these events, which happen at a frequency of 10^{-4}, depend on RecA, a protein required for homologous recombination, and can be stimulated by UV irradiation (Lin et al., 1984).

These findings on recombination events of *rhs* elements provide an indication for the acquisition and exchange of T6SS effector modules that is in some form likely common to all effector modules.

7.2.3 Evolution of auxiliary cluster 1 by stepwise acquisition and loss of effector modules

The high level of conservation within flanking regions of effector modules (chapter 4) indicates that the exchange of effector modules frequently takes place and is probably still ongoing. Here, I propose a model for the evolution of auxiliary cluster 1 as a result of the stepwise acquisition and exchange of effector modules (Figure 7-3).

In this model, I hypothesize that the original form of auxiliary cluster 1 might have been similar to the one of strain 1587 which encodes only one effector module which contains a cargo effector with antiprokaryotic activity (Figure 7-3 a). The acquisition of the actin-crosslinking domain encoded at the 3' end of *vgrG-1* is likely the result of a gene duplication event from the toxin-encoding gene *rtxA* (Sheahan et al., 2004). This acquisition could have affected the loading of the downstream encoded cargo effector onto the T6SS (Figure 7-3b). Subsequent positive selection for proper loading of the effector onto the secretion system might have adjusted the N-terminus of Tap-1 while retaining the conserved DUF4123 domain. The resulting auxiliary cluster 1 looks very similar to the one in strain AM-19226 (Figure 7-3c). The auxiliary cluster in strain V52 (Figure 7-3e) differs in the effector module encoding the cargo effector and encodes an additional open-reading frame downstream. I hypothesize that this gene encodes an orphan immunity protein that protects against a yet to be identified effector (Appendix A, Figure 9-6). I propose that acquisition of the effector module resulted in an intermediate form of the cluster with two effector modules encoding cargo effectors (Figure 7-3d). Afterwards, the 5' end of the adaptor protein-encoding gene and the effector-encoding gene of the original module might have been lost, whereas the immunity protein-encoding gene was kept (Figure 7-3e). Sana Koskiniemi and colleagues proposed a similar model for gene duplication and subsequent recombination that results in the

formation of a new *rhs* effector module (Koskiniemi et al., 2014). Whether such a mechanism applies to auxiliary cluster 1 can be tested experimentally.

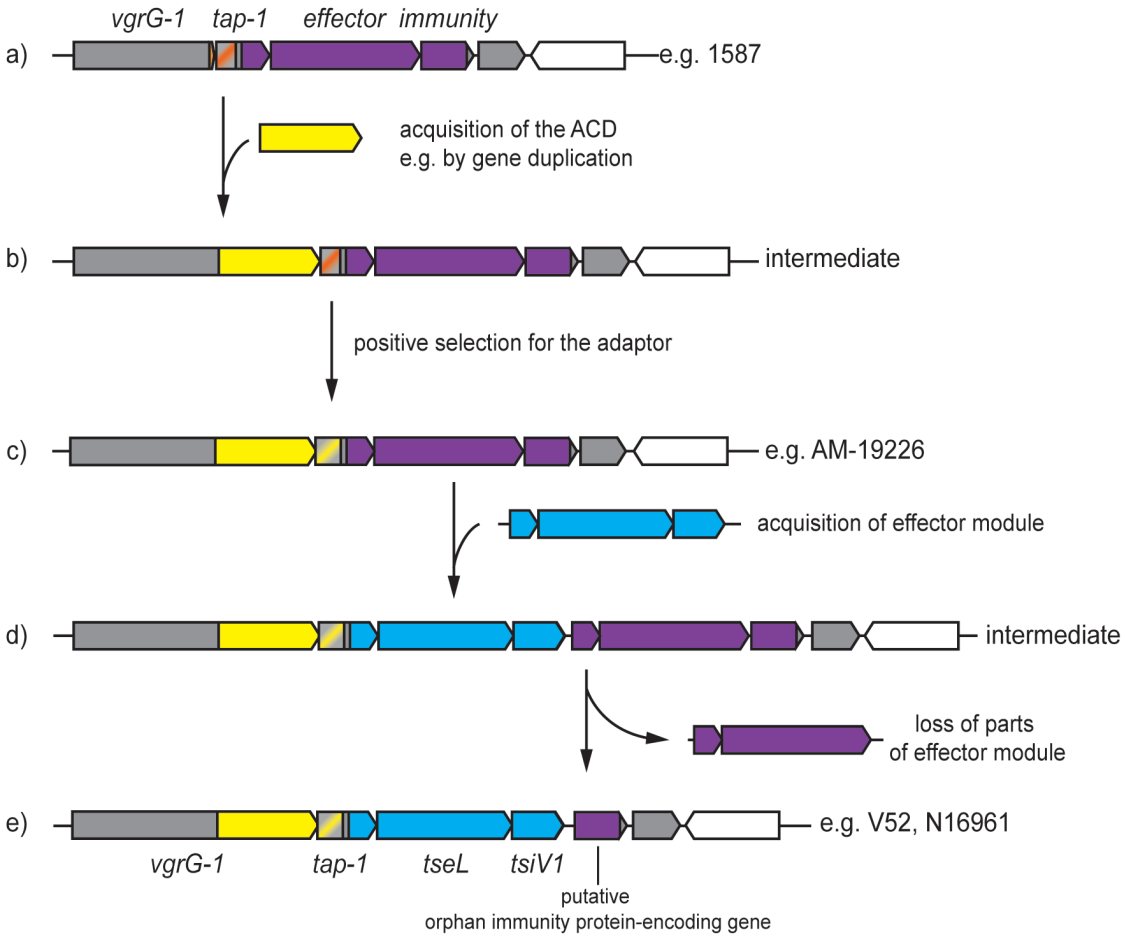


Figure 7-3 Model for the evolution auf auxiliary cluster 1.

Graphical depiction of auxiliary cluster 1 starting with *vgrG-1* and the individual modules drawn to scale. Cosegregating genes are filled with the same colour. Conserved regions are coloured in grey. Examples of strain in which a similar auxiliary cluster 1 is found are indicated on the right (see text for details).

In summary, I propose a model in which auxiliary cluster 1 evolved as the result of gene duplication, positive selection, and recombination. These events increase the number of effectors and potentially the number of immunity proteins encoded in the cluster. Additionally, the kind of effectors encoded in the cluster have changed. In light of the benefits an individual effector confers to a bacterium (see section 7.1), the

acquisition of the actin-crosslinking domain with antieukaryotic activity is an example of an event that increases the target spectrum of a bacterium to eukaryotic cells.

The remodeling of the T6SS gene cluster by exchanging effector modules poses a risk of losing other parts of the gene cluster, as the following example indicates. Effector modules could be exchanged via homologous recombination at highly conserved sites that flank effector modules. In auxiliary cluster 1, such conserved sites can be found in *vgrG-1*, in the middle of *tap-1*, and downstream of the immunity protein-encoding gene. When recombination happens in *vgrG-1* and at the 3' end of the cluster, the gain of a new effector module could result in the loss of the actin-crosslinking domain (Figure 7-4, left panel). The recombination site in the middle of *tap-1* resolves this issue by providing an internal recombination site by which effector modules can be acquired via homologous recombination without loss of the actin-crosslinking domain (Figure 7-4, right panel). In conclusion, recombination sites might be used to exchange effector modules via homologous recombination while maintaining the total number of effectors encoded in the gene cluster.

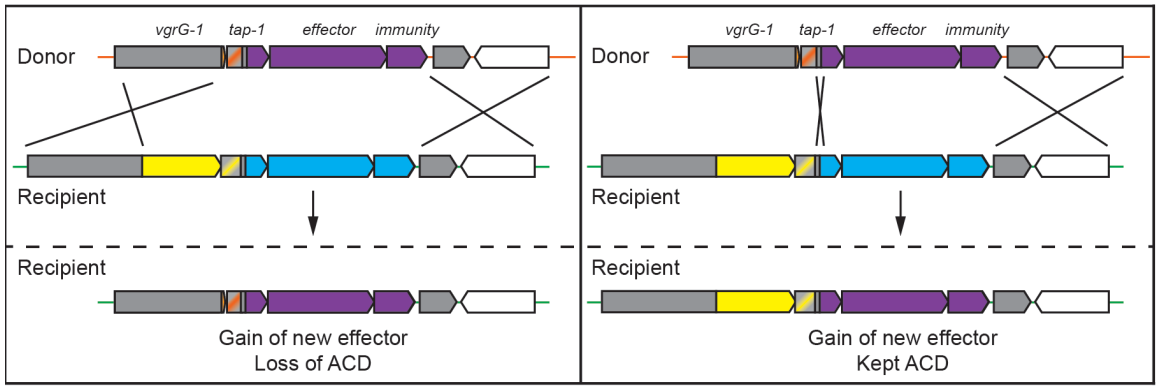


Figure 7-4 Model for the gain and loss of the ACD during remodeling of the T6SS cluster.

Graphical depiction of *vgrG-1*, *tap-1*, and effector and immunity protein-encoding genes in auxiliary cluster 1. Colour indicates differences between the nucleotide sequences, grey indicates highly conserved sequences. Crossed lines indicate sites of recombination between genomic DNA of the donor and the recipient. Two scenarios are described (left and right) that result in a different outcome for the recipient depending on the recombination site used.

7.2.4 Acquisition and distribution of effector modules requires multiple mechanisms

The mechanisms by which effector modules are acquired and exchanged needs to be elucidated experimentally. Considering the different forms of effector modules, multiple different mechanisms might be involved in such a process.

Exchange by homologous recombination is likely for effector modules encoding cargo effectors in auxiliary clusters 1 and 2, and the C-terminal extension of VgrG-3 in the large cluster. Chi sites with the sequence 5'-GCTGGTGG-3' stimulate recombination by the RecBCD pathway and might play a key role in initiating these recombination events at high frequency. Details of a mechanism that leads to the accumulation of putative orphan immunity protein-encoding genes still need to be revealed. The effector module containing *tseH* encodes a transposase that likely mediates acquisition and exchange of this module (Figure 7-2; Appendix A, Figure 9-4). Exchange of effector modules might be important for elements that contain a gene that encodes an effector required for the structural integrity of the T6SS secretion apparatus. Effector modules that contain a gene that encodes a protein that does not contribute to the structural integrity of the T6SS secretion apparatus might be acquired or lost rather than being exchanged. The original source(s) of the effector-encoding genes is unknown, also unknown are the mechanisms by which these genes were acquired, but gene acquisition is likely to have occurred by nonhomologous recombination.

7.3 The T6SS of *V. cholerae* facilitates and benefits from horizontal gene transfer

V. cholerae employs transformation, conjugation, and natural competence to acquire and exchange genomic information. Natural competence is induced during growth on chitin surfaces (Meibom et al., 2005) (Figure 7-5), which are abundant in the marine environment (Lipp et al., 2002). Chitin induces the expression of the competence regulator Tfox, which controls the expression of genes within a competence regulon of at least 19 genes (Metzger and Blokesch, 2014). Expression of the genes *comEA* and *comEC* is induced by the transcriptional regulator QstR which is activated by HapR and Tfox (Lo Scrudato and Blokesch, 2013). Natural competence allows *V. cholerae* to uptake foreign DNA, a multistep process. In the current model, a type IV pilus

translocates DNA across the outer membrane of the *V. cholerae* bacterium (Seitz and Blokesch, 2013). ComEA likely acts as a molecular ratchet and binds to the DNA that enters the periplasm (Seitz et al., 2014). The inner membrane protein ComEC transports single stranded DNA from the periplasm into the cytoplasm (Seitz and Blokesch, 2013). RecA mediates recombination between the newly acquired DNA and chromosomal DNA (Seitz and Blokesch, 2013).

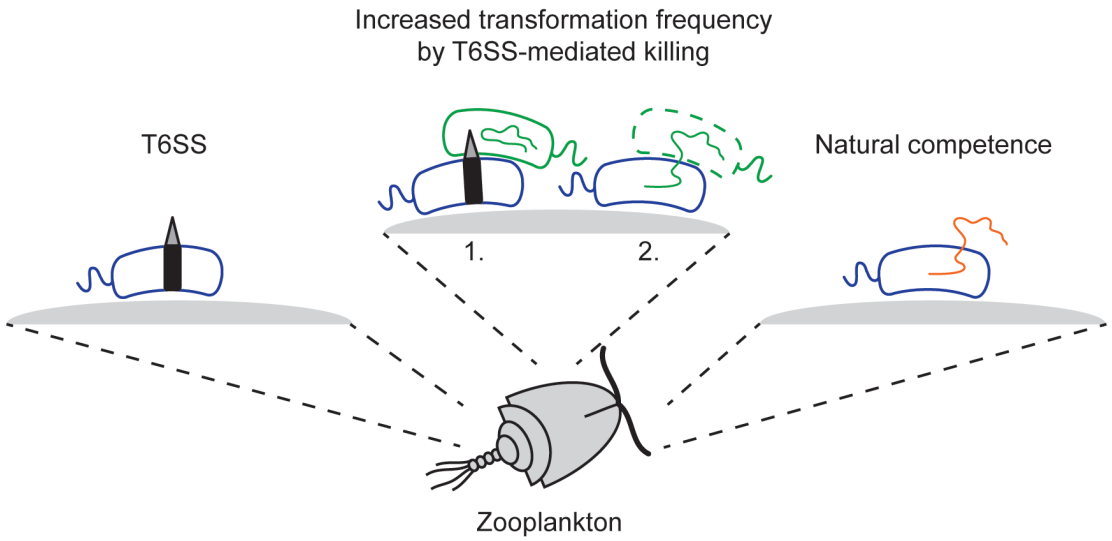


Figure 7-5 Chitin activates the T6SS and renders *V. cholerae* naturally competent.
Graphical depiction of the effects of chitin on *V. cholerae*. *V. cholerae* cells grown on chitin form an active T6SS (left). When grown on chitin, *V. cholerae* becomes naturally competent (right). In a situation in which two incompatible *V. cholerae* strains are grown on chitin, T6SS-mediated killing increases the transformation frequency (middle).

An extended analysis of genes regulated by Tfox revealed that natural competence and the T6SS are coregulated via the transcriptional regulator QstR (Figure 7-5) (Borgeaud et al., 2015). Sandrine Burgeaud and colleagues episomally expressed *tfox* in strains with an inactive T6SS under laboratory conditions and detected up to 10-fold increased expression levels of *vipA*, which encodes a protein of the outer T6SS sheath (Borgeaud et al., 2015). The authors observed extension and contraction events of the T6SS sheath when imaging *V. cholerae* with sfGFP-tagged *vipA* grown on chitin beads. Further experiments were performed to test if natural competence and the T6SS were not only genetically linked but if DNA-uptake occured upon T6SS-mediated killing

of an incompatible *V. cholerae* strain. Relocalization of ComEA, which occurs upon the uptake of foreign DNA during T6SS-mediated killing, was observed by live cell imaging to be in direct proximity to the dead bacterium. In a more quantitative approach, the authors determined the transformation frequency of a T6SS-active *V. cholerae* strain to be approximately 10^{-5}, whereas strains with a dysfunctional T6SS showed a 10-fold lower frequency. In summary, growth on chitin induces natural competence and activates the T6SS, which can be used to increase the transformation frequency by killing incompatible bacteria in close proximity to the competent bacterium.

In addition to the ability of the T6SS to promote the acquisition of foreign DNA, the T6SS has also been reported to restrict the uptake of new genetic information. Brian Ho and colleagues analyzed the transfer of the conjugative plasmid pPSV35 from *E. coli* into *P. aeruginosa* with a functional or dysfunctional T6SS and observed approximately 5-fold fewer conjugants of *P. aeruginosa* with a functional T6SS. These findings indicate that T6SS-mediated killing interferes with conjugation. The T6SS of *P. aeruginosa* differs from the T6SS of *V. cholerae* in its ability to counterattack external stimuli in a directional manner (Basler et al., 2013). To what extent the *V. cholerae* T6SS interferes with conjugation remains to be tested. T6SS-mediated killing interferes with the ability of two bacteria to get in direct contact (MacIntyre et al., 2010), which is required for conjugation. I could thus imagine a scenario, in which intraspecific competition interferes with conjugation and therefore with the exchange of conjugative plasmids and self-transmissible SXT elements (Waldor et al., 1996) between two incompatible *V. cholerae* strains. However, T6SS-mediated interference with conjugation in *V. cholerae* is limited to conditions that permit an active T6SS. We did not observe *V. cholerae* posttranslationally activating its T6SS in response to the initiation of a conjugation event, as *P. aeruginosa* does (N. Atanasova and S. Pukatzki, unpublished observation). The strain V52 with an active T6SS could thus be very efficient in preventing conjugation with incompatible bacteria. In contrast, strains like C6706 that require external stimuli to activate their T6SS might be less efficient.

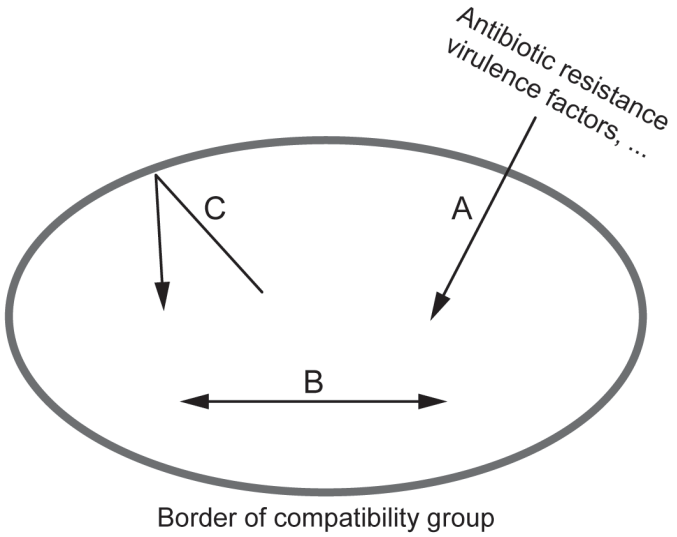


Figure 7-6 Model of gene flow directed by *V. cholerae* T6SS compatibility.

(A) Through *V. cholerae* killing of incompatible *V. cholerae* strains under conditions that induce natural competence, strains within one compatibility group gain new genomic elements like virulence factors and antibiotic resistance cassettes in the form of genomic DNA. (B) Bacteria of the same compatibility group that can exist in direct contact, distribute conjugative elements by conjugation. (C) Bacteria outside of the compatibility group of the strain with the strongest T6SS will not acquire conjugative elements or genomic DNA because they get killed when coming in direct contact with strains with a strong T6SS.

Cooperation between competence and the T6SS can affect the flow of genetic information within the *V. cholerae* species (Figure 7-6). Compatibility groups would direct the flow based on effector and immunity proteins. This model for the acquisition and distribution of mobile elements like virulence factors and antibiotic resistances is based on three hypotheses. (1) T6SS-mediated killing of a wide range of bacterial species facilitates the acquisition of foreign genes by natural competence. As tested previously, naturally competent *V. cholerae* takes up DNA of different species (Suckow et al., 2011). (2) If the acquired genes are integrated into conjugative elements, they can be distributed within bacteria of the same compatibility group in a contact-dependent manner. (3) Attempts by bacteria of other compatibility groups to acquire conjugative elements or genomic DNA would fail because incompatible bacteria are killed when they get in direct contact with incompatible bacteria for conjugation and when attempting to initiate a T6SS-mediated attack. It is possible that these events result in the uptake of less desirable

DNA, which is probably not selected for and subsequently eliminated from the population.

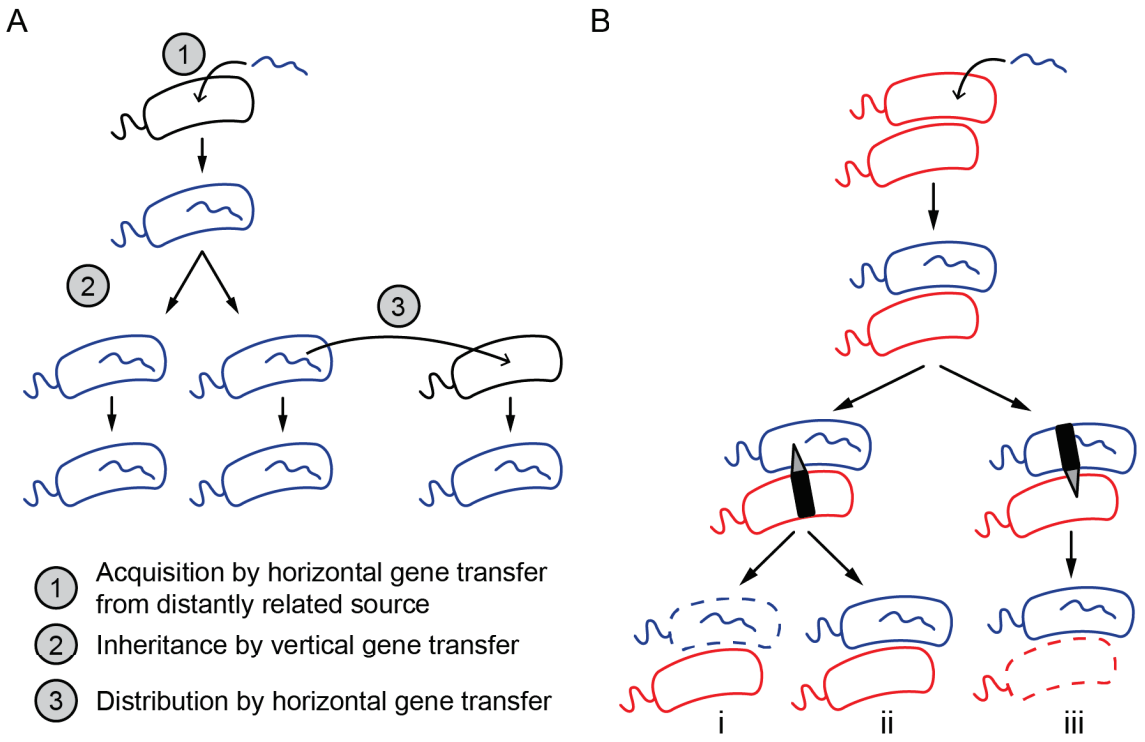


Figure 7-7 Horizontal gene transfer for the acquisition and distribution of T6SS effector modules.

(A) Horizontal gene transfer can result in the acquisition (1) and the distribution (3) of an effector module. Independent of horizontal gene transfer, a module can be inherited by vertical gene transfer (2). (B) Survival or death upon the exchange of an effector module. If exchange results in the loss of an immunity protein, the event renders the bacterium sensitive to an attack from a neighbouring bacterium (i). If the original immunity protein is maintained, the bacterium survives such an attack (ii). When attacking and killing the former kin bacterium, the bacterium that took up the effector module might also survive with the newly acquired effector (iii).

By encoding mobile gene elements in its own gene cluster, the T6SS can directly benefit from fostering horizontal gene transfer. This could be accomplished by acquiring an additional effector module (Figure 7-7 A, number 1) or by replacing a given effector module with a new module with advantageous features (Figure 7-7 A, number 3). In addition to horizontal gene transfer, T6SS effector modules might be inherited by vertical gene transfer as a result of cell division into daughter cells with genomes identical to that

of the parent cell (Figure 7-7 A, number 2). A bacterium that exchanges effector modules runs a risk of losing an immunity protein-encoding gene and therefore of losing its protection from a T6SS-mediated attack of a neighbouring cell (Figure 7-7 B, scenario i). Gaining a new effector and losing an old effector while maintaining the old immunity protein is thus an elegant way of acquiring T6SS effectors without getting killed (Figure 7-7 B, scenario ii). The collection of small open reading frames at the 5' end of auxiliary cluster 1 might be a remnant of immunity protein-encoding genes of such acquisition events (Appendix A, Figure 9-6). Alternatively, the gain of a new effector and loss of the original effector without the acquisition of an orphan immunity protein might be selected for if the new effector is superior to the effectors of the neighbouring bacteria and kills them all before the bacterium harboring the new effector is killed (Figure 7-7 B, scenario iii). I could imagine such a selection resulting in the acquisition of strong effectors that allow for intraspecific competition.

It is unclear at this point where acquisition and exchange of effector modules take place. The fact that chitin, which induces the natural competence of *V. cholerae* (Meibom et al., 2005), is found in the environment supports the notion that effector module exchange and acquisition take place in the environment. Mucin, which is abundant in the human intestine, contains N-acetylglucosamine as chitin does, and remains to be tested for its ability to induce natural competence of *V. cholerae* in vivo during infection.

The acquisition of a T6SS effector module or T6SS-independent elements or genomic DNA via horizontal gene transfer is the first step of a multistep process that contributes to bacterial evolution. As discussed in section 7.4, the acquisition of new DNA is followed by selection and multiplication events that work to maintain the DNA (Figure 7-8). Acquisition and selection can occur at the same locus or at different loci. Figure 7-7 B shows an example of acquisition of a T6SS effector module and its selection at the same locus on a chitin surface. Alternatively, a scenario can be imagined in which new DNA is acquired under conditions of natural competence in the environment and the new trait is selected for during subsequent infection of a host. The observation that pandemic strains contain the same T6SS effector modules might indicate that these modules were selected for during infection. From where and from what these modules were acquired is not yet known.

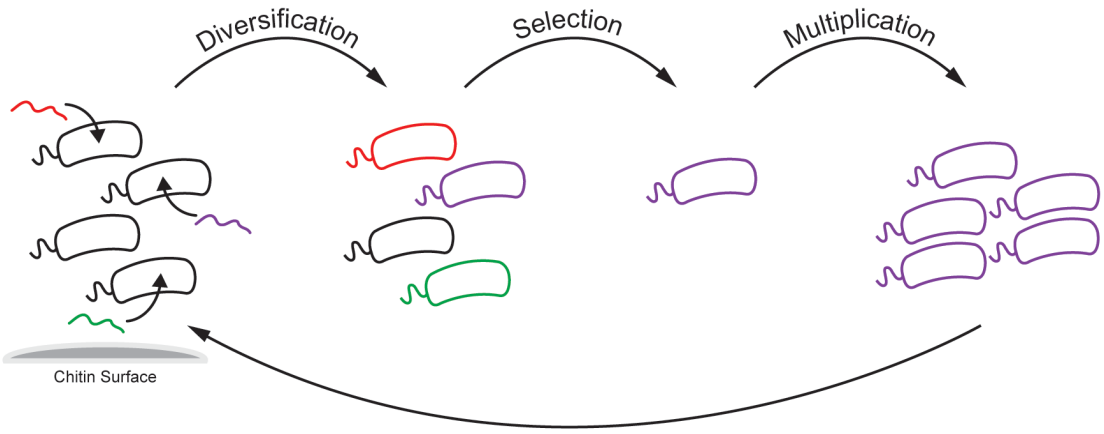


Figure 7-8 Model of bacterial evolution upon acquisition of new DNA.

Graphical depiction of a bacterial population that diversifies by acquiring various pieces of DNA. In a situation of selective pressure, one piece of DNA confers an advantage to a bacterium compared to the bacteria that acquired different pieces of DNA. Subsequent multiplication of the advantaged bacterium establishes the new trait in the population. This population might enter the cycle again under conditions that induce the uptake of foreign DNA.

Taken together, T6SS-mediated killing facilitates the uptake of genomic DNA by natural competence. Increased horizontal gene transfer might benefit the whole bacterium or only the T6SS, which could acquire new effector modules.

7.4.1 The role of intraspecific competition in *V. cholerae* as a species

We and others have noticed that bacterial strains that belong to the same compatibility group share common features and are closely related (Hill et al., 1995; Unterweger et al., 2014). I revisit this observation and discuss two scenarios, one in which the acquisition of T6SS effector modules drives the evolution of a species and another in which T6SS effector modules are acquired as a result of diversification.

7.4.1 T6SS effectors and the structure of a species

When comparing the phylogeny of *V. cholerae* strains, based on six house-keeping genes as representatives of the core genome, to the T6SS effector modules they

encode, we made two observations (Unterweger et al., 2014). The *V. cholera* strains Amazonia and TM11079-80 belong to the same compatibility group. Their close relationship is indicated in the phylogenetic tree in which these strains are located in the same vertical line as their most recent common ancestor (MRCA) (Unterweger et al., 2014). In the same phylogenetic tree, 12129(1) and LMA3984-4 are located at the end of separate horizontal lines that are derived from the same MRCA (Unterweger et al., 2014), indicating that the two strains are derived from the same MRCA but are now more distantly related. The strains 12129(1) and LMA3984-4 belong to different compatibility groups (Unterweger et al., 2014). An example of a group of strains that originated from the same common ancestor is the group of pandemic strains that belongs to the compatibility group AAA (Unterweger et al., 2014).

Differences in T6SS effector module sets between strains of the same species have been described for multiple species. Examples for species other than *V. cholerae* are *Serratia marcescens* (Murdoch et al., 2011) (hospital-acquired infections), *Pseudomonas aeruginosa* (Jiang et al., 2014) (pulmonary infections), and *Proteus mirabilis* (Gibbs et al., 2008) (kidney stones). Bioinformatics analyses by Harrison and colleagues suggests intraspecific differences of the T6SS among *Campylobacteri jejuni* isolates (Harrison et al., 2014). Hill and colleagues compared the phylogeny of *E. coli* strains to the *rhs* loci before it was known that they encode T6SS effectors (Hill et al., 1995). This analysis revealed a correlation between clonal groups of the species and the *rhs* elements they encode.

I speculate that the link between T6SS effector modules and the phylogeny of a species is the result of diversification and selection events in which T6SS-mediated killing might play a role in some cases but does not need to always play a major role (see section 7.4.2 for more specific scenarios).

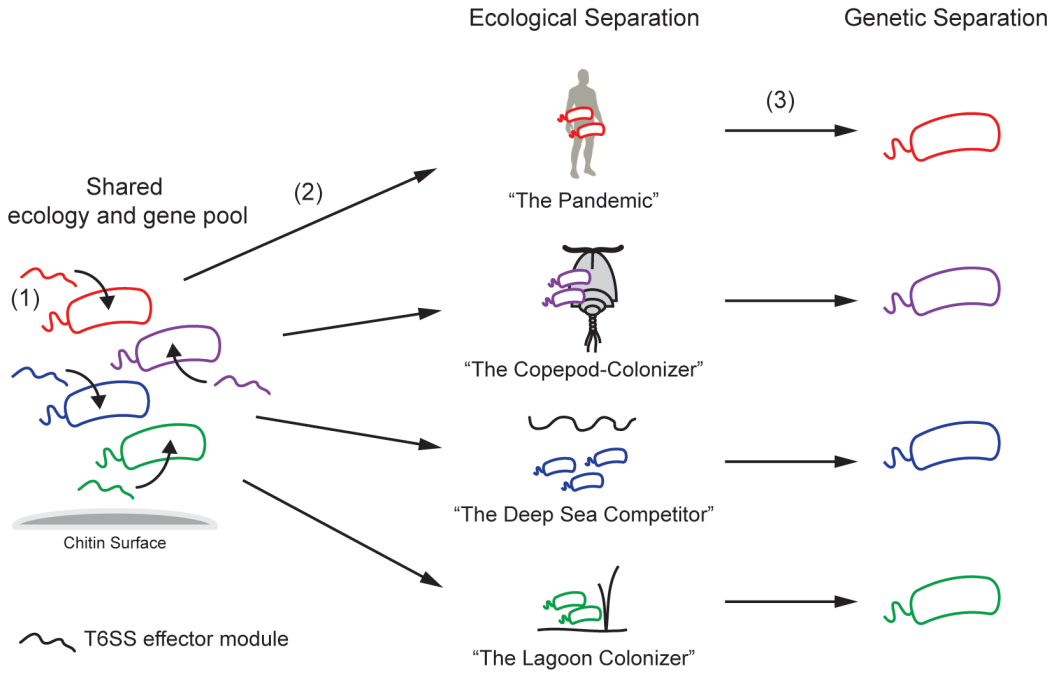


Figure 7-9 Model of diversification of a species as a result of ecological and genetic separation.

(1, 2) Diversification by acquisition of effector modules (indicated by colour change of the bacteria) results in the ability to colonize different ecological niches. (3) Separate acquisition of mutations and other genetic elements in the core genome results in genetic separation.

7.4.2 T6SS effector modules as the driver of species diversification

T6SS-mediated diversification and selection events could lead to the ecological separation of bacteria within a species (Figure 7-9). Bacteria likely face different competitors unique for the niche they colonize. The T6SS might be a tool for bacteria to combat prokaryotic and eukaryotic competitors in the respective niche. Individual T6SS effectors differ in their target range and the condition in which they are toxic. Subsequently, niches with the same competitors could select for bacterial strains with the same T6SS effector modules. This form of selection could result in genetically different strains that share the same T6SS effector module set. Strains that colonize different niches are likely to encounter different competitors and are expected to differ in their T6SS effector modules. Ecological separation resulting in genetic separation could

explain the observation that two distantly related strains derived from the same common ancestor differ in their T6SS effector module (Unterweger et al., 2014). Strains that colonize separate niches could acquire mutations and take up additional mobile elements independent from each other. Subsequent phylogenetic analysis of the core genome would locate strains in separate clades on individual branches of a tree.

7.4.3 T6SS effector module acquisition as the result of diversification

Alternatively, *V. cholerae* strains could also differentiate independent of the T6SS and colonize a variety of ecological niches. As a result of the new environment in the new niche, they might take up niche-specific T6SS effector modules (Figure 7-10). These events would also lead to a correlation between the phylogeny of strains and the T6SS effector module set they harbor.

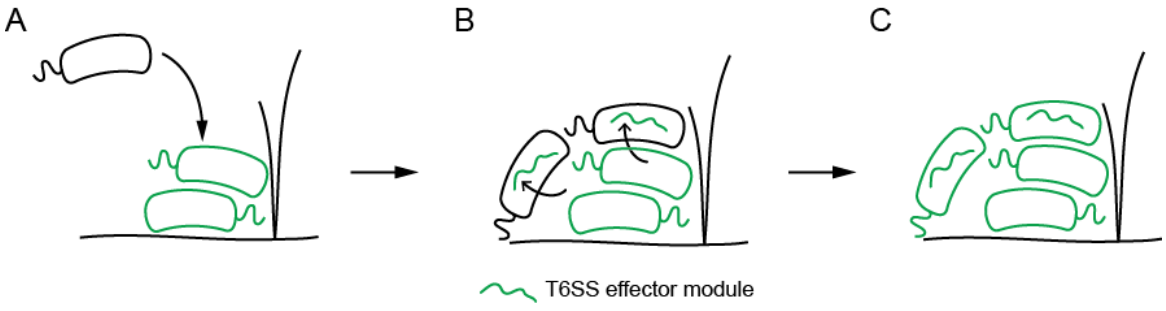


Figure 7-10 Acquisition of an effector module upon colonization of a niche.
(A) *Vibrio cholerae* colonizes a new niche colonized by other bacteria. (B) Bacteria colonize and, as a result of genetic exchange with local bacterial communities, acquire effector modules from the local *V. cholerae* population. (C) Bacteria in one niche share the same effector module.

An example of the influence of the local environment on the gene flow of integrons, which are also mobile elements, is provided in a study on *V. cholerae* in the USA and Bagladesh (Boucher et al., 2011). Yan Boucher and colleagues compared *V. cholerae* strains from Cape Cod (USA) with other *Vibrio* species isolated from the lagoon on the Cape and *V. cholerae* strains isolated in Bangladesh. The core genome of *V. cholerae* isolates from Cape Cod turned out to be closer related to *V. cholerae* strains from Bangladesh than to *Vibrio metoecus* (Kirchberger et al., 2014) from Cape Cod. In

contrast, the integrons of *V. cholerae* from Cape Cod were more similar to *V. metecus* isolated from the Cape Cod than to *V. cholerae* from Bangladesh. Integrons consist of an integrase and an *attI* site (Hall and Collis, 1995). The integrase inserts genes flanked by an *attC* site by mediating site-specific recombination between the *attC* and the *attI* site. Multiple genes, called cassettes, can be integrated into integrons. A promoter downstream of the integrase-encoding gene drives expression of the cassettes. A superintegron is found on the chromosome of *V. cholerae* that contains over 100 cassettes (Mazel et al., 1998). Most of the proteins encoded in the superintegron remain to be characterized in regard to their function. Some of the proteins encoded in the superintegron confer antibiotic resistances (Rowe-Magnus and Mazel, 1999). This study by Yan Boucher and colleagues suggests that the exchange of mobile elements across species borders occurs frequently (Boucher et al., 2011). The T6SS effector modules in this sample set remain to be analyzed.

7.4.4 The T6SS and the evolution of *V. cholerae* as a pathogen

Species like *V. cholerae* in which only a subset of strains is pathogenic raises the question of how the T6SS contributes to the evolution of pathogenic and pandemic *V. cholerae*. The T6SS itself is found in all strains sequenced to date independent of their source of isolation. The effector set common to all pandemic strains might have been acquired following the above described principle of diversification and selection. The human host provides a unique niche with different competitors and selection pressures than the environment. The observation that already a 2nd pandemic strain (isolated in 1849) belonged to the AAA compatibility group indicates an acquisition of the AAA effector module set during the evolution of pandemic strains before 1849. Since this acquisition event, the effector module set is likely inherited because this set is found in all pandemic strains irrespective of date or place of isolation. The conservation of the effector module set over at least 150 years might indicate that it is selected for and that acquisition of other effector modules might have conferred a selective disadvantage. The observation that individual effector modules of the pandemic strains are also found in distantly related, environmental isolates indicates that clinical and environmental strains

and environmental strains engage in genetic exchange (Figure 7-11). For example, the environmental isolate of the non-O1/non-O139 serogroup strain VL426 encodes the same effector module in the large cluster as the pandemic strains (Unterweger et al., 2014). If the exchange of an effector module requires direct contact between donor and recipient, and if the environmental isolate of the nonO1/O139 serogroup strain VL426 acquired its effector module in the large cluster directly from a pandemic strain, the pandemic strain and the environmental strain must have gotten in contact somewhere and exchanged the effector module (Figure 7-11). The life cycle of pandemic *V. cholerae* that involves the environment as a reservoir would support this model of a niche overlap, the possibility of exchange of genomic DNA, and also the colonization of different niches.

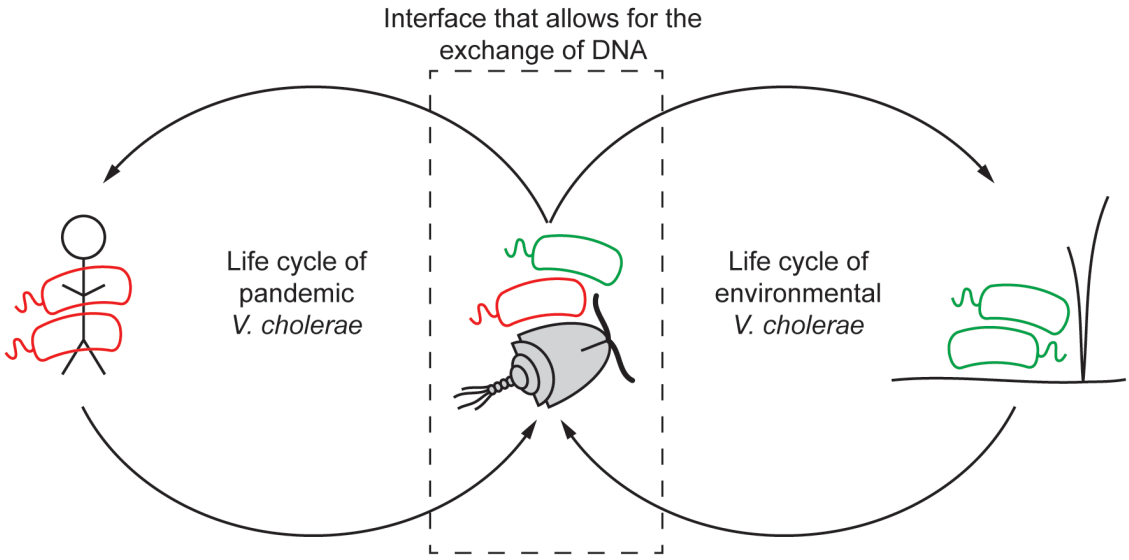


Figure 7-11 Model of the exchange of effector modules between *V. cholerae* strains.
Graphical depiction of the life cycle of pandemic and environmental *V. cholerae*. Overlap of the two life cycles is indicated, that could explain the exchange of effector modules between pandemic and environmental *V. cholerae* strains.

7.5 The role of the T6SS and intraspecific competition of *V. cholerae* as a pathogen

Intestinal pathogens face multiple challenges when colonizing the small intestine including the immune system, members of the microbiota, and other bacteria of the same

species. Here, I describe putative roles of the T6SS in allowing *V. cholerae* to overcome these challenges and establish a successful infection.

7.5.1 The T6SS and its interactions with the immune system

One role of the T6SS might be to counteract the immune response of the human body during infection. The immune system presents a burden to infection by intestinal pathogens. Macrophages are part of the innate immune response and among the first cells to respond to foreign antigens. One of the T6SS effectors is VgrG-1 with a C-terminal extension that causes actin-crosslinking of murine macrophages (Pukatzki et al., 2007). The presence of this effector in all pandemic strains but not in all nonpandemic strains suggests that pandemic strains with an ACD are selected for (see chapter 5). Amy Ma incubated infant mice for 16–18 h with high titres of wild-type *V. cholerae* or a mutant lacking the ACD and recovered 1,000-fold less CFUs of the ACD-deficient mutant than wild-type (Ma and Mekalanos, 2010). When inoculating the mice with the same titre only for 2 hours, no difference in the recovered CFUs could be determined. These findings indicate that the ACD-deficient mutant does not have a colonization defect per se but rather lacks the ability to maintain colonization at high titres. These results were further supported in an experiment in which V52ΔACD was administered to mice that were precolonized with wild-type V52 or V52ΔACD (Ma and Mekalanos, 2010). 1,000-fold more V52ΔACD was recovered from mice precolonized with V52 compared to the mice precolonized with V52ΔACD (Ma and Mekalanos, 2010). These data indicate that V52ΔACD does not have a colonization defect per se but that a proportion of the bacterial population needs the ACD – possibly to protect the whole population from being diminished by the immune system. Other parts of the study also revealed a proinflammatory effect of the ACD as determined by neutrophil infiltration in the gut and fluid accumulation. I speculate that this is one of the reasons why pandemic *V. cholerae* strains tightly regulate their T6SS and use external cues like mucin and bile acids to activate and deactivate the T6SS in response to their location and time during infection. These experiments illustrate that the T6SS harbors an antieukaryotic effector that is immunogenic itself but is also required to maintain colonization.

A homologous region to the ACD in VgrG-1, which mediates antieukaryotic activity, is found in the enterotoxin A (*rtxA*) encoded in a gene adjacent to *ctx* on the large chromosome of *V. cholerae*. Both domains were shown to crosslink actin in eukaryotic cells (Sheahan et al., 2004). It will be interesting to further determine the advantage of encoding this domain in a T6SS effector versus the advantage of a toxin that is secreted by the type I secretion system (Boardman and Satchell, 2004). One advantage might be that the T6SS directly translocates the effector into the target cell whereas the type I secretion system secretes an effector into the extracellular space only.

7.5.2 The T6SS as a tool to fight off the microbiota

Another function of the T6SS might be to counteract colonization resistance mediated by an intact microbiota. Bacteria of the microbiota are known to interfere with the infection of intestinal pathogens by interference with their quorum sensing (Duan and March, 2010; Hsiao et al., 2014), competition for nutrients (Kamada et al., 2012), and antibacterial activity (Russell et al., 2014b).

Another limiting factor for *V. cholera*e infection might be the limited access to nutrients (Vogt et al., 2015). An experiment with two intestinal pathogens other than *V. cholerae* gives us an idea of how competition for nutrients can affect infection of the intestine. Kamada and colleagues colonized germ-free mice with *C. rodentium*, which solely depends on monosaccharides, and *B. thetaitomicron,* which can metabolize monosaccharides and polysaccharides (Kamada et al., 2012). When the mice were fed a conventional diet of mono- and polysaccharides, over 10^{11} CFUs of *C. rodentium* were recovered per gram of feces. Only about 10^{9} CFUs of *C. rodentium* were recovered when the mice were fed monosaccharides that both species compete for. These results indicate that dependence on the same nutrient source can be a limiting factor during colonization and multiplication in the gut. To test if *V. cholerae* outcompetes Gram-negative bacteria of the microbiota during infection in a T6SS-dependent manner, I suggest two experiments. (1) Gram-negative members of the microbiota could be systematically tested for their sensitivity to T6SS-mediated killing in vitro. (2) An alternative experimental approach would be to analyse human microbiota in germ-free mice upon

colonization with either wild-type *V. cholerae* or a T6SS-deficient mutant by deep sequencing.

Another hurdle for bacterial pathogens to overcome are antibacterial mechanisms the microbiota employs to actively fight invaders. Bioinformatics analyses of the genomes of human microbiota revealed widespread biosynthetic gene clusters that potentially produce antibiotic compounds (Donia et al., 2014). The T6SS itself also has been found among species of the gut microbiota, for example, in *Bacteroides fragilis*. Russell and colleagues detected expression of the T6SS gene *tssC* in samples from the cecum of germ-free mice monocolonized with *B. fragilis* (Russell et al., 2014b). Growth competition experiments performed in vitro revealed significantly fewer *B. thetaitomicron*, a Gram-negative member of the microbiota, in the presence of *B. fragilis* with a functional T6SS than in the presence of *B. fragilis* with a dysfunctional T6SS, indicating that *B. fragilis* has a T6SS with antiprokaryotic properties. This observation is also of interest to our understanding of the *V. cholerae* T6SS during infection. Recent findings indicate that the products of bile acid metabolism by commensal bacteria like *B. thetaitomicron* downregulate T6SS activity of *V. cholerae* (V. Bachman, B. Kostiuk, and S. Pukatzki, unpublished observation). How *V. cholerae* responds to a T6SS-mediated attack of bacteria of the microbiota like *B. fragilis* needs to be further investigated. I suggest that killing assays be performed between T6SS-active species of the microbiota and *V. cholerae* in an anaerobic environment to test if *V. cholerae* is able to resist these attacks and is able to outcompete the attacking bacteria.

In summary, these examples show that the microbiota provide multiple hurdles for *V. cholerae* to overcome in order to colonize and multiply in the intestine. The ability to kill certain members of the microbiota via effectors of the AAA compatibility group might provide a benefit during colonization and multiplication. To specifically test the role of the AAA effector set, the ability of wild-type C6706 and C6706 in which the AAA effector set is replaced by a different effector set, to colonize infant mice could be compared in single infection and competition experiments.

7.5.3 T6SS-mediated intraspecific competition to select for toxin-producing strains

The T6SS might also have a function as a mechanism for intraspecific competition between *V. cholerae* strains. The T6SS might be used by virulence factor-producing strains to kill *V. cholerae* strains that do not produce virulence factors (Figure 7-12). The main virulence factor of *V. cholerae* is cholera toxin. Cholera toxin induces water efflux from intestinal epithelial cells. Although this water efflux causes dehydration of the patient, it can be beneficial to intestinal pathogens that can use the water efflux to get back in the environment. Among strains of the species *V. cholerae* that coexist in the environment, only a subgroup of strains contains the cholera toxin-encoding genes. Early in an outbreak, before clonal expansion of the virulent strain, humans likely take up a mixed inoculum of toxigenic and nontoxigenic strains *V. cholerae*. That is, *V. cholerae* strains that produce and that do not produce cholera toxin. Whereas all of the strains would use up nutrients and resources, only some would take the burden of contributing to escape the host again by producing cholera toxin. However, there might be a threshold for the level of cholera toxin that must be produced to induce water efflux. T6SS-mediated intraspecific competition might be a mechanism for cholera-toxin producing strains to kill nonproducing strains and use the required nutrients to multiply sufficiently to reach a level of cholera toxin production that induces water efflux. Isolates of multiple clonal complexes were prepared from the environment and the stool of cholera patients during local outbreaks in Bangladesh (Rashed et al., 2014; Stine et al., 2008). When genotyping the individual isolates based on variable-length tandem repeats at 5 loci (VC0147, VC0437, VC1650, VCA0171, and VCA0283), the isolates could be grouped into clonal complexes (Rashed et al., 2014). Ten isolates from six patients were further analyzed (Rashed et al., 2014). In one patient, all 10 isolates belonged to the same genotype. Another patient had isolates from two unrelated genotypes. The isolates of the remaining eight patients belonged to various genotypes that were all related (Rashed et al., 2014). The isolates remain to be tested for their compatibility group membership.

Whereas T6SS-mediated intraspecific competition of *V. cholerae* in vivo still needs to be tested, a study from the Mekalanos group indicated that T6SS-mediated competition takes place in vivo between two bacteria of the same strain (Fu et al., 2013). When inoculating an infant rabbit with wild-type C6706 and C6706Δ*vipA*Δ*tsiV3*, which

lacks the cognate immunity protein to VgrG-3 and a functional T6SS to not get killed by an isogenic sister cell, Fu and colleagues recovered 10-fold less CFUs of the immunity mutant than of wild-type from the intestine after an 18 hour-long incubation (Fu et al., 2013). These observations are further supported by similar results from infant mice (Figure 9-10) and provide a proof of principle that T6SS-mediated killing occurs in vivo.

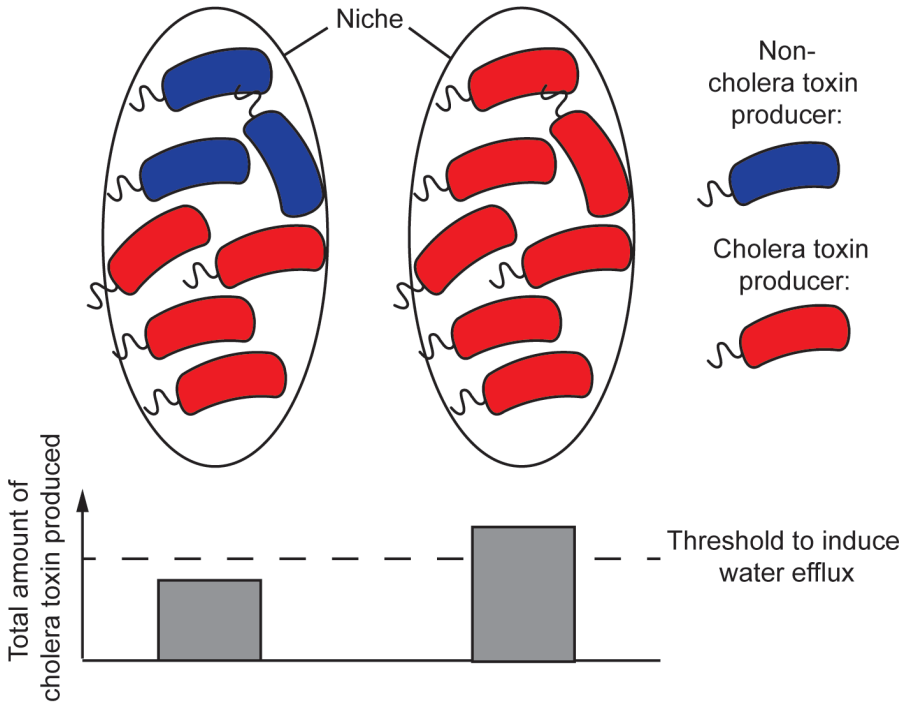


Figure 7-12 Model of the role of intraspecific competition to reach the threshold of cholera toxin production.

Graphical depiction of a niche in the intestine that provides limited resources of nutrients and space. This niche is either colonized by a mix of cholera-toxin producing and nonproducing bacteria (left) or by cholera-toxin producing bacteria only because the nonproducing bacteria got killed in a T6SS-dependent manner. At the bottom, a hypothetical graph of the total amount of cholera toxin produced in the individual scenarios is shown.

The observation that all analyzed pandemic strains belong to the AAA compatibility group whereas most of the nonpandemic strains do not supports our model of the T6SS as a mechanism to select for toxin-producing strains. Strains like M66-2 that belong to the AAA compatibility group but do not encode cholera toxin can be considered as cheaters that do not contribute to the cost of producing cholera toxin but

are able to coexist with and benefit from cholera-toxin producing strains during infection (Unterweger et al., 2014).

Classical biotype strains lack the ability to engage in T6SS-mediated killing (chapter 6). The ability of classical biotype strains to produce higher levels of cholera toxin than El Tor biotype strains (Kaper et al., 1995) might explain how classical strains could successfully cause pandemics without making use of the T6SS. If in a mixed inoculum with nontoxigenic strains, the classical biotype strains might produce cholera toxin above the threshold to induce watery diarrhoea without the need to kill nontoxigenic strains.

Taken together, T6SS-mediated intraspecific competition might be a mechanism for the selection of toxin-producing strains within a mixed inoculum of *V. cholerae* strains (Figure 7-12).

7.5.4 T6SS and its potential role in clonal populations during infection

Recent advances in characterizing the first steps of intestinal colonization further exemplify the potential role of the T6SS in *V. cholerae* (Abel et al., 2015).

Abel and colleagues inoculated an infant rabbit with a barcoded population of *V. cholerae* and compared the total population size to the number of barcodes within the population at 13 sites of the intestine. Even though the population size was very similar 20 h post infection, the size of the founding population differed between locations of the intestine (Abel et al., 2015). Further studies showed that the size of founding populations varies at individual sites over the time course of infection. These findings suggest that colonization of the intestine is a dynamic process marked by clonal expansion of bacteria that are exposed to varying bottlenecks in dependency on the site of colonization and time-point during infection.

Another recent study focused on the initial colonization of the intestine by *V. cholerae*. Y. Millet and colleagues inoculated infant mice with an equal ratio of green and red fluorescently labeled *V. cholerae* bacteria (Millet et al., 2014). Imaging of the intestine showed either red or green microcolonies of bacteria in the crypts and at the base of villi in the small intestine. These results suggest that individual *V. cholerae* cells

colonize the small intestine and subsequently develop clonal microcolonies. The number of microcolonies then likely constitutes the founding population observed in the analysis described above.

If surrounded by bacteria of other species or the same species, the T6SS might provide a mechanism for bacteria to differentiate self from nonself based on the compatibility group they belong to.

7.6 The concept of compatibility groups

We have started to develop a concept that subdivides *V. cholerae* strains into compatibility groups based on their ability to coexist or outcompete each other. This is the draft of a concept that has to be further developed, considering T6SS-independent killing between strains (Unterweger et al., 2012), potential orphan immunity proteins, additional effectors that will be identified in the future, and potentially differential expression of T6SS effectors. We can already see that this concept has implications for our classification of bacteria and the potential for applications as biomarker and a novel therapeutic strategy.

Compatibility groups provide the basis to further investigate the role of T6SS-mediated killing on the biology of the species. We already have indications that T6SS effector module sets might impact the phylogenetic structure of a bacterial species. For the analysis of a bacterial species, it might be increasingly important to consider the ability of strains to coexist in direct contact with each other and to compete with other bacteria.

T6SS effector modules could serve as biomarkers. Because of the correlation between the virulence potential of a strain and the compatibility group it belongs to, a killing assay to test for coexistence with an AAA strain might be a simple test that does not require electricity and could be used to monitor *V. cholerae* in the environment to prevent outbreaks, for water management during outbreaks, and for diagnostics of clinical samples. A similar concept, called bacteriocin typing, has been previously proposed (Israil et al., 1983). The advances in genome sequencing and the increased

awareness of the importance of bacteria-bacteria interactions for a bacterial species might help to promote the concept of T6SS effector modules as biomarkers.

Compatibility groups could also be used as the basis for two separate therapeutic approaches. Pandemic strains of the AAA compatibility group could specifically be targeted without affecting other strains of the species. One strategy could be to let pandemic *V. cholerae* bacteria kill each other; normally they do not do so because they belong to the same compatibility group (Laura Diaz-Satizabal and Stefan Pukatzki, unpublished). Killing of compatible bacteria could be achieved by interfering with immunity proteins that protect from a T6SS-mediated attack of bacteria of the same compatibility group. The introduction of cheaters could be another strategy. These cheaters would be in the same compatibility group as pandemic strains without harboring the virulence factors. Cheaters like the vaccine strain Peru-15 would use the resources of pandemic strains during infection and might not allow pandemic strains to grow to high enough densities to reach the threshold of cholera toxin production to cause diarrhoea.

The ability of the T6SS to load a variety of effectors could also be used to engineer strains with a particular effector set that can be used as weapons against pathogens. I could imagine a scenario in which a bacterial species of the microbiota is equipped with a T6SS to which pathogens are sensitive to. In summary, the concept of compatibility groups forms the basis of a variety of potential applications. The relevance of intraspecific competition for *V. cholerae* has to be further investigated to differentiate between the importance of a variety of effectors to kill bacterial competitors of other species and the use of these diverse effectors to kill other *V. cholerae* strains.

7.7 Overall summary of conclusions

Our findings on T6SS-mediated intraspecific competition of *V. cholerae* contributes to a better understanding of the T6SS. The discovery of a variety of effectors in *V. cholerae* strains exemplifies the diversity of T6SS effector proteins and their enzymatic activities. VgrG proteins play a key role in translocating effectors outside the bacterium. Some effectors depend on adaptor proteins like Tap-1 to be loaded onto the secretion system.

Comparison of T6SS gene clusters of multiple *V. cholerae* strains revealed some genes to be conserved whereas others to be highly variable. Genes encoding structural and regulatory proteins are conserved whereas genes encoding effector and immunity proteins are part of effector modules that are likely to be exchanged between *V. cholerae* strains by horizontal gene transfer.

The observation of closely related strains with the same T6SS effector module set and distantly related strains with different effector module sets suggests a putative role of the T6SS during the diversification of the species. All analyzed pandemic *V. cholerae* strains belong to the same compatibility group characterized by the same set of effector modules and the ability to coexist among each other in direct contact.

Although likely not required for strains to cause past pandemics, the T6SS might play multiple roles for strains of the current 7th pandemic to mediate interactions with the immune system, with bacteria of the microbiota, and with other strains of the same species. Intraspecific competition might be a tool for pandemic strains to outcompete nonpandemic strains that do not contribute to the cost of producing virulence factors and take up limited resources like space and nutrients. The characteristic T6SS effectors of pandemic strains might also be beneficial to outcompete bacterial species abundant in the human microbiota and rarely encountered in the environment.

These findings contribute to our basic understanding of the bacterium *Vibrio cholerae* and the type VI secretion system and have the potential to be developed into multiple applications for the benefit of human health.

Chapter 8
Materials and Methods

8 Materials and Methods

Strains and culture conditions

Bacteria were grown in liquid culture in Luria-Bertani broth (Maniatis, 1987) (1% Tryptone, 0.5% Yeast extract, 0.5% NaCl) on the roller drum (full speed) at 37ºC. Alternatively, bacteria were grown on LB agar plates (LB broth, 1.5% agar). If necessary, liquid broth or LB plates were supplemented with the following antibiotics at the indicated concentrations: Streptomycin (100µg/ml), Rifampicin (50µg/ml), Kanamycin (50µg/ml), Ampicillin (100µg/ml). Antibiotic stock solutions were dissolved in ddH_2O or DMSO (for Rifampicin).

Bacteria were stored in cryovials in 20% glycerol at -80 ºC. A complete strain list is provided at the end of this chapter.

Dictyostelium discoideum AX3 cells were maintained in liquid culture in HL5 media (per liter: 10g glucose, 5g yeast extract, 5g proteose peptone, 5g thiotone peptone, 0.67g Na_2HPO_4, 0.34g KH_2PO_4) with shaking at 22ºC. The culture was derived from the Dicty stock centre (Chicago, IL; http://dictybase.org/StockCenter/StockCenter.html).

Primer design

The nucleotide sequence of the template was derived from NCBI (http://www.ncbi.nlm.nih.gov/). Primers were designed manually to ideally end at the 3' end with repetitive Gs or Cs and to melt at a temperature of at least 60ºC. The webtool OligoAnalyzer (https://www.idtdna.com/calc/analyzer) was used to determine the melting temperature and to test for the likelihood of primer-dimers. Primers ideally form homodimers with a delta G value of less than 10% of their maximal delta G. If required, restriction sites were added to the 5' end of primers. Because I used blunt end ligation following the initial amplification, no additional nucleotides were added at the 5'end. If required, FLAG (GAC TAC AAG GAC GAC GAT GAC AAG)- or HIS (CAT CAT CAC CAT CAC CAC)-tags were added. If long primers with additional sequence non-homologous to the template were created, they were created in a way that at least half of the primer was binding to the original template. Primers were ordered from the company

Integrated DNA Technologies (IDT; Coralville, IA, USA) or Eurofins MWG Operon (Huntsville, AL, USA).

Polymerase chain reaction (PCR)

PCR reactions of 25µl or 50µl were prepared. The smaller volume was prepared for reactions in which PCR was used to control for the presence of an insert, for example. The larger reaction volume was used for reactions in which the amplified DNA was used for subsequent cloning steps. Reactions were mixed in dependency on the polymerase used. TopTaq polymerase (Qiagen, Toronto, ON, Canada) without proofreading activity was used for PCR reactions in which the product was not used for subsequent cloning whereas phusion polymerase (ThermoScientific, Waltham, MA, USA) was used for reactions in which the product was used for subsequent cloning. For amplifications from plasmids, 1µl of purified plasmid was used. For amplifications from genomic DNA, 3.5µl of a 1/10 dilution of overnight culture in ddH_2O was used. Polymerase, template, primers, dNTPs and buffers were used as recommended by the vendors of the polymerases. 34 amplification cycles were run in the thermo cycler. The extension time was adjusted to the polymerase used. A melting temperature of 60ºC was chosen.

Agarose gel electrophoresis

To separate the amplified PCR fragments by length, I performed agarose gel electrophoresis. 1% agarose gels were prepared and Red Safe (1/20,000; iNtRON Biotechnology, Korea) added to later visualize the DNA of the gel. If still required, loading dye was added to the PCR products and the whole solution was loaded onto the agarose gel. Gels were run for 35min at 120V. A geldoc system (Cell Biosciences) was used to visualize the DNA bands on the gel and take pictures.

Gel extraction

To gel extract PCR fragments of an indicated size, I used a gel extraction kit (ThermoScientific, Waltham, MA, USA). The piece of agarose gel containing the band of interest was dissolved in 700µl resuspension solution and loaded onto the columns to concentrate the DNA on the filter. After 1 washing step with 700µl wash buffer, DNA

was eluted with 30µl of preheated ddH_2O. Yields were determined using the NanoDrop (ThermoScientific, Waltham, MA, USA). Individual steps were performed as suggested by the vendor.

Blunt-end ligation into pJet

I usually cloned PCR products directly into the plasmid pJet using the pJet easy cloning kit (ThermoScientific, Waltham, MA, USA). Purified PCR-template was used as a template and dNTPs, ligase and blunt-end digested pJet was added. The reaction was incubated at room temperature for 30min and afterwards transformed into electrocompetent *E. coli* TOP10 cells (see detailed description electroporation).

Preparation of electrocompetent cells

Slightly differing protocols were used to make small or large batches of electrocompetent cells. For small batches (Gonzales et al., 2013), a needle pin size loop full of bacteria from an overnight culture of bacteria was resuspended in cold ddH_2O for *E. coli* (or cold 2mM $CaCl_2$ for *V. cholerae*) and washed two times using a pre-cooled centrifuge. Afterwards, the cells were resuspended in 40µl of ddH_2O (or 2mM $CaCl_2$) and directly used for electroporation.

Transformations via electroporation

For electroporation, I usually added 3µl (range from 1-7µl) from the ligation reaction or purified plasmid to the competent cells. The mixture was added onto cuvettes (2mm wide) and electroporated at 2.5V and 200 Ohms. Afterwards, the cell suspension was resuspended in 1ml of LB and incubated for 1 hour at 37ºC on the roller drum before plating out 100µl and all remaining cells onto LB agar plates supplemented with the required antibiotic to select for transformants.

Plasmid purification

Plasmids were purified from bacterial cultures using the miniprep kit from ThermoScientific (Waltham, MA, USA) or FroogaBio (Toronto, ON, Canada). 3-5ml of bacterial culture were resuspended in resuspension buffer. Lysis buffer and neutralization

buffer were added, followed by a centrifugation step. Supernatant was loaded onto a column, washed with washing buffer and eluted with 30µl of ddH_2O. I followed the instructions of the vendor.

Restriction digest

I used fast digest restriction enzymes (Life Technologies; now ThermoFisher Scientific, Waltham, MA, USA) to extract inserts flanked by restriction sites from plasmids. Reactions of 1µg were set up for digests to control for the presence of an insert, 3µg of DNA were digested if the insert was used for a following ligation. In the example of the digest of 1µg DNA, the respective amount of plasmid was mixed with 1µl of restriction enzyme each, 2µl of FD buffer and filled up to a final volume of 20µl with ddH_2O. The reaction was incubated at 37ºC for 30min. If a vector was digested prior to ligation, intestinal calf phosphatase (Invitrogen; now ThermoFisher Scientific, Waltham, MA, USA) was added to the reaction and incubated for 5min at 37ºC following the ligation. Agarose gel electrophoresis and subsequent gel extraction was used to get purified insert.

Sitcky end ligation

To ligate inserts into the vector pBAD24, vector and insert were digested (for details see restriction digest). The products were ligated using the T4 ligase (Thermo Scientific, Waltham, MA, USA) according to the vendor's instructions. Vector and insert were mixed at a molar ratio of 1:3 to a final total volume of 100ng. The reaction was incubated at room-temperature for one hour, followed by electroporation into *E. coli* TOP10 except for ligations into the suicide plasmid pWM91 that were transformed into *E. coli* DH5alpha lambda pir.

Cloning of genes into vetctors of the pBAD series

I amplified genes by PCR with primers flanked with single cutter restriction sites that cut in the multiple cloning site of the pBAD vector and not within the insert. For some vectors of the pBAD series, additional sequence was added to provide the Kozak sequence. The genes were cloned into pJet, their presence validated by restriction digest

and then ligated into the pBAD vector and transformed into *E. coli* TOP10. The presence of the insert was validated by restriction digest. If necessary, the plasmid was purified and transformed into the respective *V. cholerae* strain of interest.

Cloning of constructs for the BACTH assay

To test the direct interaction between T6SS proteins (see BACTH assay for detailed description), genes were cloned into the plasmids pKET25, pUT18 and pUT18-C kindly provided by the Raivio lab. Procedures described above were used. The genes were cloned with or without start or stop codons to allow translational fusions of the gene to the coding sequence of the T25 or T18 fragment. The multiple cloning sites of the respective vectors were used.

Cloning of constructs for gene knock-out and knock-in

To knock-in or knock-out genes by allelic replacement (Metcalf et al., 1996), a construct was created that contains overlapping regions for recombination and either a scar that does not contain the gene of interest any more or contains additional sequence for a knock-in. The designated construct was cloned and ligated into pWM91. Restriction sites were chosen that do not cut within the knock-out construct.

To generate the construct, 6 primers A-F were designed. Forward primer A binds approximately 800bp upstream of the gene of interest and contains a restriction site. Reverse primer D binds approximately 800bp downstream of the gene of interest and contains a different restriction site. To generate a scar to delete the gene of interest, primer C consists of a sequence complimentary to the 4 nucleotides upstream the gene of interest, followed by the first 15 nucleotides at the 5' and the last 15 nucleotides at the 3' end of the gene of interest, and ends with the 4 nucleotides downstream of the gene of interest. Primer B is a reverse complement of primer C. Primers E and F are used to screen for successful knock-outs. The forward primer E binds upstream of primer A, the reverse primer F binds downstream of primer D.

The 4 primers A-D were used for a two-step PCR reaction for which the strain of interest is used as template to generate the construct for interest. In the first step, two separate reactions with primers A, B and C, D were performed. In a second step the

products of the PCR reactions of step 1 were used as a template for a reaction with primers A and D. The resulting construct was ligated into pWM91. Subsequently, this vector was transformed into *E. coli* SM10λpir.

Mating and allelic exchange

First, pWM91 is mated into the strain of interest, followed by two rounds of selection, one selection for a cross-in of pWM91 into the chromosome of the strain of interest and a second round of selection for a cross-out of the plasmid and the original sequence, leaving the construct originally inserted in pWM91 behind (Metcalf et al., 1996).

To generate a knock-out, the respective strain and *E. coli* SM10λpir harboring the respective knock-out construct are grown overnight in a lawn on LB agar plates supplemented with the required antibiotic. The bacteria are harvested and distributed onto a 4 by 4 cm^2 area onto a fresh, pre-warmed LB agar plate. After incubation for about 8h, the bacteria are harvested, serially diluted and plated onto LB agar plates supplemented with the selective antibiotics for the strain of interest and pWM91 to select for transformants. The transformants were picked the next day and re-streaked one more time onto selective plates. Colonies were then picked, inoculated into 1ml of LB supplemented with antibiotics and grown for 4 hours on the roller drum. Afterwards, the bacteria are serially diluted and plated onto sucrose plates (1% tryptone, 0.5% yeast, 6% sucrose, 2% agar). The plates are incubated at room temperature for 2 days. Colonies are picked and screened by PCR with primers E and F to confirm the loss of the gene and plated onto LB plates supplemented with the antibiotic resistance of the plasmid to make sure that the strain is sensitive to this bacterium as a result of plasmid loss.

The scar and surrounding regions were sequenced to make sure that the scar is in frame and no point-mutations were introduced in surrounding genes.

Bacterial killing assay

The protocol of the bacterial killing assay originates from Dana MacIntyre and colleagues (MacIntyre et al., 2010) and is used to test if two bacteria kill each other in a T6SS-dependent manner during a 4h-long co-incubation period on an LB agar plate.

Bacteria are grown over night on selective LB agar plates. To mix a specific number of bacteria with each other the next day, bacteria were harvested of the plate and resuspended in 2ml LB. A 1/50 dilution of this resuspension was used to measure the OD600 and calculate the amount of bacteria per ml resuspension. Depending on the experiment, predator and prey were mixed at a 10:1 or 1:1 ratio. When combined at a 10:1 ratio, 10^8 and 10^7 bacteria were mixed. When combined at a 1:1 ratio, 10^8 were mixed with 10^8 bacteria. A test-tube was prepared in which the number of bacteria for 5 killing spots was resuspended in 125µl of LB. 25µl of this resuspension was spotted onto an LB agar plate. This plate contains 20ml of LB agar and was incubated at room-temperature over-night and pre-warmed for 30min at 37ºC. If the killing assay requires expression of genes from a plasmid, L-arabinose was added to the plate to a final concentration of 0.1% if not indicated otherwise. To determine the number of bacteria per killing spot based on a CFU count, bacteria were taken out of the test tube with the resuspension for 5 killing spots, serially diluted and plated onto LB agar plates supplemented with antibiotics to either select for predator or prey. The killing plate was incubated at 37ºC for four hours. After 4h, the killing spot was harvested and resuspended in 1ml of LB followed by a serial dilution and plating onto LB plates supplemented with antibiotics to select for predator or prey. The following day, CFUs were counted and the number of bacteria per killing spot before and after the 4h killing spot determined. If indicated, the competitive index was calculated by dividing the ratio of predator and prey after 4 h killing with the ratio of predator and prey before the 4h killing period.

Killing can vary based on dryness of the killing plate. In my experience especially VasX-mediated killing and to a lesser degree VgrG-3-mediated or TseL-mediated killing are affected by dryness of the killing plate.

To test if a protein of interest was expressed during the 4h killing period, 200µl were taken out of the test tube in which the killing spot was resuspended after 4h of killing, and the bacteria therein resuspended in 100µl of sample buffer.

High-throughput killing assays

To screen the Keio library (Baba et al., 2006) for T6SS-resistant mutants of *E. coli*, a modified version of the killing assay was developed that could be used at high throughput. Bacterial cultures were grown in 96 well-plates over-night, diluted 1/100 the next day and grown to an OD600 of 0.5. Each *E. coli* mutant was mixed with either *V. cholerae* V52 or *V. cholerae* V52Δ*vasK* at a ratio 1:1 and 2µl were spotted onto a 96 well-plate filled with 150µl of LB agar. The plates were incubated at 37ºC for four hours. Afterwards, the bacteria were washed of the LB agar and transferred into a new 96 well-plate. LB supplemented with Kanamycin that selects for the growth of *E. coli* was added. The OD600 of the cultures was determined and the plate incubated at 37ºC with shaking for another 6 hours. The OD600 was determined and the difference between the OD600 at these two time-points calculated.

Plaque assay

A plaque assay was performed to determine the virulence of *V. cholerae* against amoeba. Therefore, 100µl of bacterial overnight culture and 10^3 *D. discoideum* cells were plated onto SM/5 plates (2g glucose, 2g bactopeptone, 0.2g yeast extract, 0.1g $MgSO_4$, 1.9g KH_2PO_4, 1g K_2HPO_4, 2.92g NaCl, 7.5g agar in 1l ddH_2O, pH 6.5). The plates were incubated at room temperature for three days and photographed.

Protein secretion profile

Bacterial cultures of the pellet and supernatant are harvested and further concentrated to test for the expression and secretion of T6SS proteins.

To detect TseL in culture supernatants, overnight culture was diluted 1/100 into a culture with a total volume of 5ml and grown in the presence of a selective antibiotic to an OD600 of ~0.5 at 37ºC on the roller drum. If required, arabinose was added to the culture to a final concentration of 0.1% and the culture was grown for an additional hour.

The OD600 was determined and 1200µl were transferred into a test tube and centrifuged at 4ºC with 14,000 rpm for 10min. The tubes were kept on ice and the supernatant was transferred to a new test tube after filtration through a 0.22µm filter. 250µl of trichloracetic acid was added and the supernatant samples were stored stationary and over-night rotating at 4ºC. The pellet was resuspended in a volume of sample buffer calculated by multiplying the volume of culture taken (µl) with the OD600 and dividing the resulting number by 3. The next day, the supernatant samples were centrifuged for 1 h at 14,000rpm followed by two washes in ice-cold acetone. The pellets were air-dried and resuspended in 30µl sample buffer (10% glycerol , 0.08M Tris-HCl (pH 6.8), 2% SDS, 0.01% bromophenol blue, 1.25% β-mercatptoeathanol).

To analyze the secretion of Hcp, mid-log cultures with an OD600 of 0.5 were used.

SDS-polyacrylamide gel electrophoresis (SDS PAGE)

Cell lysates were subjected to SDS PAGE to separate proteins by size for later detection via western-blotting.

In dependency on the size of the proteins of interest, 10 to 15% SDS polyacrylamide gels were used. The separating gel was made of acrylamide, 0.37M Tris pH8.8, 3.4mM SDS, 2mM ammonium persulfate (APS) and 0.6µM tetramethylethylenediamine (TEMED). The stacking gel was made of acrylamide, 0.12M Tris pH6.8, 3.4mM SDS, 3mM APS and 1.4mM TEMED. 10-20µl of sample was loaded onto a gel with 10 wells. Page ruler (Thermo Scientific, Waltham, MA, USA) was used as a size marker. Gels were run in the presence of Laemmli running buffer (25mM Tris, 0.2M glycine, 0.1% SDS) at 100V and maximal Amp for 30min and then 150Vand maximal Amp until the samples separated throughout the gel.

Western-blotting

Proteins separated by size were transferred onto a nitrocellulose membrane and specific antibodies were added to detect the protein of interest.

Transfer of the proteins from the polyacrylamide gel onto the nitrocellulose membrane (BioRad) was performed using the liquid transfer apparatus (BioRad) at

175mAmp for 90 minutes. The membrane was then incubated in blocking buffer (0.1M Tris, 1M NaCl, 0.5% Tween-20, pH 8.0, supplemented with 5% skim milk powder) for 1 hour at room temperature. The following antibodies were used at the indicated titres for incubation over 1 hour at room-temperature or overnight at 4ºC: rabbit anti-Hcp (1:500, (Pukatzki et al., 2006)), rabbit anti-TseL (1:5000, Southern Alberta Cancer Institute, Calgary, Canada), mouse anti-DnaK (1:15000, Enzo Life Sciences, Farmingdale, NY, USA), mouse anti-HIS (1:1000, Santa Cruz, Dallas, TX, USA) and mouse anti-FLAG (1:1000, Sigma, St. Louis, MO, USA). After three washes with TBST, the membrane was incubated for 1 hour at room temperature with the following secondary antibodies linked to horse-radish peroxidase: goat anti-rabbit (1:3000, Santa Cruz, Dallas, TX, USA), goat anti-mouse (1:3000, Santa Cruz, Dallas, TX, USA), anti-mouse native protein (1:1000, Abcam, Toronto, ON, Canada).

Substrate (SuperSignal West Pico, Thermo Scientific, Waltham, MA, USA) was added to the peroxidase and Fuji medical X-ray film was used to detect light emitted by the peroxidase bound to antibodies.

Immunoprecipitation

Immunoprecipitation experiments were performed to test the interaction between T6SS proteins. Therefore, bacterial lysates were incubated with beads coupled to FLAG-antibodies. The beads were separated from the remainder of the culture and protein complexes bound to the FLAG antibody were analyzed by SDS-PAGE and western-blotting.

Overnight cultures were inoculated 1/100 into a total of 250 ml of LB supplemented with antibiotics and grown at 37°C shaking to an OD of 0.5. Arabinose was added to a final concentration of 0.1% and cultures were grown for an additional hour. Afterwards, bacteria were resuspended in 10 ml of TBST buffer (50 mM Tris, 150mM NaCl, 0.05% Tween 20, pH 7.6) supplemented with protease inhibitor cocktail (cOmplete Tablets EDTA-free, Roche, Basel, Switzerland) and 1mg/ml lysozyme (Sigma, St. Louis, MO, USA). The resuspended cells were lysed by sonication (3 bursts for 30 sec, level 3) and centrifuged at 10,000g for 15min at 4°C. The supernatant was transferred to a falcon tube and kept on ice. Part of the lysate was mixed with sample

buffer, boiled and later loaded as "total". The remaining part of the lysate was used for the immunoprecipitation.

To prepare the anti-FLAG beads, 40 µl of gel suspension (A2220, Sigma, St. Louis, MO, USA) per sample was added to a test tube with ddH_2O and rotated for 5min at 4°C. Beads were centrifuged at 5,000g for 30 sec, settled for 1min and washed two more times with 1ml of TBST. The suspension was distributed equally among the samples. The samples were then rotating for 2h at 4°C. Afterwards, the beads were washed with TBST, resuspended in 20µl sample buffer and boiled. These samples were further processed in SDS-PAGE and western-blot.

Bacterial two-hybrid assay (BACTH assay)

Bacterial two hybrid assays were performed to test the direct interaction between T6SS proteins. Therefore, liquid bacterial cultures were grown overnight in the presence of selected antibiotics. 5µl culture was spotted onto LB agar plates with a volume of 20ml supplemented with 40µg/ml X-Gal, 0.5mM IPTG, Ampicillin and Kanamycin. The plates were incubated at 30ºC until bacteria turned blue except for the negative control that remained white. Pictures of the plates were taken using an htc phone.

Animal experiments

To determine if T6SS-mediated killing happens *in vivo*, we conducted animal experiments using the suckling mouse model for cholera (Klose, 2000).

5-day old mice were inoculated intragastrically with a 1:1 mix of a *lacZ* positive and *lacZ*-negative strain at a total number of 10^9 bacteria/50µl. Serial dilutions of the input mix were prepared to determine the exact number of bacteria by CFU count. The mice were kept over-night and sacrificed to remove the small intestine. The tissue was resuspended in PBS and homogenized. Serial dilutions of the homogenate were prepared and plated onto LB plates supplemented with streptomycin and X-Gal (40µg/ml) to determine the ratio of bacteria that colonized the small intestine. The experiment was approved by the respective Ethics Committee at the University of Alberta.

Bioinformatics

To predict the three-dimensional structure of a protein, the webtool Phyre2 (Kelley et al., 2015) was used. For the modeling of TsiV1 and TseL in chapter 6, the intensive mode was used.

Alignments between nucleotide or amino acid sequences of individual proteins or genes were preformed using the standard settings of MUSCLE (Edgar, 2004) in the Geneious software (Kearse et al., 2012). Alignments of nucleotide sequences that encode more than one open-reading frame were performed using the progressive MAUVE algorithm (Darling et al., 2004). The nucleotide or amino acid identity between sequences and the GC content of a particular nucleotide sequence was determined using the Geneious software.

Multiple methods have been used to build phylogenetic trees. The phylogenetic tree of 37 *V. cholerae* strains shown in chapter 4 was assembled based on the concatenated nucleotide sequences of *adk* (adenylate kinase, locus tag VC0986), *gyrB* (DNA gyrase subunit, locus tag VC0015), *mdh* (malate dehydrogenase, locus tag VC0432), *recA* (recombinase A, locus tag VC0543), *pgi* (glucose-6-phosphate isomerase, locus tag VC0374) and *rpoB* (DNA-dependent RNA polymerase, locus tag VC0328) to establish the relatedness of *V. cholerae* to the corresponding genes from *V. mimicus* as the outgroup (Boucher et al., 2011). The data set containing these homologues had 38 sequences and 11,046 positions. Alignments were done using ClustalW (Larkin et al., 2007). All positions were determined to be homologous and were included in the alignment. jModelTest2 (Darriba et al., 2012; Guindon and Gascuel, 2003) was used to find the best model of nucleotide evolution for the sequences, incorporating corrections for invariable sites as well as a four-category gamma correction for rate variation when appropriate. A substitution model that accounted for both invariant sites and gamma-distributed rate heterogeneity (GTR,I,G) was chosen. PhyML v. 2.4.4 (Guindon and Gascuel, 2003) was used for maximum likelihood analysis, and to generate ML bootstrap values based on 100 pseudoreplicates of each data set. The phylogenetic tree shown is the PhyML consensus topology. All other trees shown in the results section were created using the standard setting of RaxML (Stamatakis, 2006) with 100 bootstraps.

Statistical analysis

Un-paired parametric t-tests were calculated based on log-transformed data from killing assays to determine the statistical significance of differences between two indicated groups of data points. The F-test was used to get an indication for equal variance between the compared groups. If necessary, we used the Welch correction to account for non-equal variance. P-values smaller than 0.05 were considered as statistically significant.

Software and links

For the bioinformatics analyses, the software Geneious (version 8, (Kearse et al., 2012)) was used. The program figtree (http://tree.bio.ed.ac.uk/software/figtree/) was used to display phylogenetic trees. The following web tools were used for cloning: NEB cutter (http://nc2.neb.com/NEBcutter2/), Oligoanalyzer (https://www.idtdna.com/calc/analyzer) and Reverse complement converter (http://www.bioinformatics.org/sms/rev_comp.html). The software Prism was used to create graphs. Adobe CC was used to make graphical depictions. Microsoft Office was used for multiple applications. EndNote was used to manage the references.

Table of bacterial strains

Strain	Description	Reference and Source
V. cholerae V52	O37 serogroup strain isolated in Sudan in 1968 from a clinical sample; Streptomycin resistant	(Zinnaka and Carpenter, 1972), John Mekalanos (Harvard Medical School, Boston, MA, USA)
V. cholerae V52 *vgrG-1::myc*	V52 with endogenously myc-tagged *vgrG-1*	(Pukatzki et al., 2007)
V. cholerae V52Δ*vgrG-1*	In frame deletion mutant lacking *VCV52_1390*	This study
V. cholerae V52Δ*vgrG-3*	In frame deletion mutant lacking *VCV52_A0139*	(Brooks et al., 2013)
V. cholerae V52Δ*tap-1*	In frame deletion mutant lacking *VCV52_1391*	This study
V. cholerae V52Δ*tseL*	In frame deletion mutant lacking *VCV52_1392*	(Unterweger et al., 2014)
V. cholerae V52Δ*vasX*	In frame deletion mutant lacking *VCV52_A0043*	(Miyata et al., 2013)
V. cholerae V52Δ*tseL*Δ*vasX*	In frame deletion mutant lacking *VCV52_1392, VCV52_A0043*	(Miyata et al., 2013)
V. cholerae V52Δ*tseL*Δ*vgrG-3*	In frame deletion mutant lacking *VCV52_1392, VCV52_A0139*	(Miyata et al., 2013)
V. cholerae V52Δ*vasX*Δ*vgrG-3*	In frame deletion mutant lacking *VCV52_A0043, VCV52_A0139*	(Miyata et al., 2013)
V. cholerae V52Δ*tseL*Δ*vasX*Δ*vgrG-3*	In frame deletion mutant lacking *VCV52_1392, VCV52_A0043, VCV52_A0139*	(Miyata et al., 2013)
V. cholerae C6706	O1 serogroup strain (biotype: El Tor, serotype: Inaba) isolated in Peru in 1991 from a clincal sample; resistant to Streptomycin and Rifampicin	(Thelin and Taylor, 1996), John Mekalanos (Harvard Medical School, Boston, MA, USA)
V. cholerae C6706Δ*tseL*	In frame deletion mutant of	(Miyata et al., 2013)

	tseL	
V. cholerae C6706Δ*tsiV1*	In frame deletion mutant of *tsiV1*	(Miyata et al., 2013)
V. cholerae C6706Δ*vasX*	In frame deletion mutant of *vasX*	(Miyata et al., 2013)
V. cholerae C6706Δ*tsiV2*	In frame deletion mutant of *tsiV2*	(Miyata et al., 2013)
V. cholerae C6706Δ*vgrG-3*	In frame deletion mutant of *vgrG-3*	(Miyata et al., 2013)
V. cholerae C6706Δ*tsiV3*	In frame deletion mutant of *tsiV3*	(Miyata et al., 2013)
V. cholerae C6706Δ*tsiV1*Δ*tsiV2*	In frame deletion mutant of *tsiV1* and *tsiV2*	(Miyata et al., 2013)
V. cholerae C6706Δ*tsiV1*Δ*tsiV3*	In frame deletion mutant of *tsiV1* and *tsiV3*	(Miyata et al., 2013)
V. cholerae C6706Δ*tsiV2*Δ*tsiV3*	In frame deletion mutant of *tsiV2* and *tsiV3*	(Miyata et al., 2013)
V. cholerae C6706Δ*tsiV1*Δ*tsiV2*Δ*tsiV3*	In frame deletion mutant of *tsiV1, tsiV2* and *tsiV3*	(Miyata et al., 2013)
V. cholerae 1587	Clinical isolate from Peru, isolated 1994, O12 serogroup strain	(Chun et al., 2009), Michelle Dziejman, (University of Rochester, New York)
V. cholerae 1587Δ*tap-1*	In frame deletion mutant of A55_1501	This study
V. cholerae 1587Δ*A55_1501-1503*	In frame deletion mutant of A55_1501-1503	This study
V. cholerae V52ΔHRH	V52 lacking *hapA*, *rtxA*, *hlyA*	(Raskin et al., 2006)
V. cholerae V52ΔHRHΔ*vasX*	V52 mutant lacking *vasX*	(Miyata et al., 2011)
V. cholerae V52ΔHRHΔ*vasW*	V52 mutant lacking *vasW*	(Miyata et al., 2013)
V. cholerae V51	O141 serogroup strain, isolated in the USA in 1987 form a clinical sample	Michelle Dziejman (Univeristy of Rochester, New York)
V. cholerae 1587	O12 serorgoup strain, isolated in Peru 1994 from	Michelle Dziejman (Univeristy of Rochester,

	a clinical sample	New York)
V. cholerae DL4211	O123 serogroup strain, isolated from the Rio Grande river delta (USA) in 2008	(Unterweger et al., 2012), Daniele Provenzano (UBrownsville, TX, USA)
V. cholerae DL4211Δ*vasK*	In frame deletion mutant of *vasK*	(Unterweger et al., 2012)
V. cholerae MZO-2	O14 serogroup strain isolated in Bangladeh in 2001 from a clinical sample	Jun (Jay) Zu (University of Pennsylvania)
V. cholerae AM-19226	Non-O1/O139 serogroup strain isolated in Bangladesh in 2001 from a clinical sample	(Dziejman et al., 2005), John Mekalanos (Harvard Medical School, Boston, MA, USA)
V. cholerae DL4215	O113 serogroup strain isolated from the Rio Grande river delta (USA) in 2008 from an environmental sample	(Unterweger et al., 2012) Daniele Provenzano (UBrownsville, TX, USA)
V. cholerae DL4215Δ*vasK*	In frame deletion mutatn of *vasK*	(Unterweger et al., 2012)
V. cholerae MZO-3	O37 serogroup strain isolated from Bangladesh in 2001 from a clinical sample	Michelle Dziejman (University of Rochester, New York)
V. cholerae 2740-80	O1 serogroup strain (biotype: El Tor, serotype: Inaba) isolated at the Gulf Coast (USA) in 1980 from an environmental sample	Jun (Jay) Zu (University of Pennsylvania)
V. cholerae MAK757	O1 serogroup strain (biotype: El Tor) isolated in Indonesia in 1937 from a clinical sample	Michelle Dziejman (University of Rochester, New York)
V. cholerae MO10	O139 serogroup strain isolated in India in 1992 from a clinical sample	Michelle Dziejman (University of Rochester, New York)
V. cholerae N16961	O1 serogroup strain (biotype: El Tor, serotype: Inaba) isolated in Bangladesh in 1975 from a clinical sample	John Mekalanos (Harvard Medical School, Boston, MA, USA)

V. cholerae CA401	O1 serogroup strain (biotype: classical) isolated in India in 1953 from a clinical sample	Shelley Paine (University of Texas at Austin)
V. cholerae O395	O1 serogroup strain (biotype: classical) isolated in India in 1965 from a clinical sample	John Mekalanos (Harvard Medical School, Boston, MA, USA)
V. cholerae NIH41	O1 serogroup strain (biotype: classical) isolated in India in 1940	John Mekalanos (Harvard Medical School, Boston, MA, USA)
V. cholerae DL2111	Strain isolated from the Rio Grande river delta (USA) in 2008 from an environmental sample	(Unterweger et al., 2012), Daniele Provenzano (UBrownsville, TX, USA)
V. cholerae DL2112	Strain isolated from the Rio Grande river delta (USA) in 2008 from an environmental sample	(Unterweger et al., 2012), Daniele Provenzano (UBrownsville, TX, USA)
P. phenolica	Isolated from the Rio Grande river delta (USA) in 2008 from an environmental sample	(Unterweger et al., 2012), Daniele Provenzano (UBrownsville, TX, USA)
V. communis	Isolated from the Rio Grande river delta (USA) in 2008 from an environmental sample	(Unterweger et al., 2012), Daniele Provenzano (UBrownsville, TX, USA)
V. harbeyi	Isolated from the Rio Grande river delta (USA) in 2008 from an environmental sample	(Unterweger et al., 2012), Daniele Provenzano (UBrownsville, TX, USA)
E. coli MG1655	F- lambda- ilvG- rfb-50 rph-1, Rifampicin-resistant	(Blattner et al., 1997)
E. coli TOP10	F- mcrA Δ(mrr-hsdRMS-mcrBC) φ80lacZΔTΔM15 lacX74 nupG recA1 araD139 Δ(ara-leu_7697 galE15 galK16 rpsL(Str^R) end Ai λ^-	Invitrogen
E. coli DH5alph λpir	fhuA2 D(argF-lacZ)U169 phoA glnV44 W80 D(lacZ)M15 gyrA96	(Elliott and Kaper, 1997)

	recA1 relA1 endA1 thi-1 hsdR17	
E. coli SM10 λpir	KmR, thi-1, thr, leu, tonA, lacY, supE, recA::RP4-2-Tc::Mu, pir	(R. Simon, 1983)

Table of plasmids

Plasmid	Description	Reference
pBAD24	pBAD vector, pBR332 ori, araC, Amp^R	(Guzman et al., 1995)
p*tsiV1*	pBAD24 carrying *tsiV1* (N16961)	(Miyata et al., 2013)
p*tsiV1*^{163F}	pBAD24 carrying *tsiV1* (CA401)	This study
p*tsiV2*	pBAD24 carrying *tsiV2* (N16961)	(Miyata et al., 2013)
p*tsiV3*	pBAD24 carrying *tsiV3* (N16961)	(Miyata et al., 2013)
p*vgrG-1*	pBAD24 carrying *VCV52_1390*	(Pukatzki et al., 2007)
p*tap-1*^{V52}	pBAD24 carrying *VCV52_1391*, C-terminally HIS tagged	This study
p*tap-1*$^{AM-19226}$	pBAD24 carrying *A33_1345*, C-terminally HIS tagged	This study
p*tap-1*^{1587}	pBAD24 carrying *A55_1501*, C-terminally HIS tagged	This study
p*1503::FLAG*	pBAD24 carrying *A55_1503,* C-terminally FLAG tagged	This study
p*tap-1*^{V52}*::FLAG*	pBAD24 carrying FLAG tagged *VCV52_1391*	This study
pFLAG::*tap-1*^{V52} (AA1-286)	pBAD24 carrying FLAG tagged *VCV52_1391*	This study
pFLAG::*tap-1*^{V52} (AA15-286)	pBAD24 carrying FLAG tagged *VCV52_1391* lacking the first 42 basepairs	This study
p*vgrG-1::HIS*	pBAD24 carrying HIS tagged *vgrG-1*	(Pukatzki et al., 2007)
p*vgrG-1::FLAG*	pBAD24 carrying FLAG tagged *vgrG-1*	(Pukatzki et al., 2007)
p*vgrG-1* AA1-1163	pBAD24 carrying *VCV52_1390*	This study
p*vgrG-1* AA1-679	pBAD24 carrying *VCV52_1390,* basepairs 1-2037	This study
p*vgrG-1* AA1-642	pBAD24 carrying *VCV52_1390*, basepairs 1-1926	This study

Table of primers

Name	Sequence (5' to 3', restriction sites are underlined)	Description
tsiV1-A	GGATCC GCCATAGCTTAGGGGGCGC	Primer A, knock-out *tsiV1* in C6706
tsiV1-B	GGCATTAATTATCATCAGAATTCAATA ACTTCATCTTATTTGC	Primer B, knock-out *tsiV1* in C6706
tsiV1-C	TAAGATGAAGTTATTGAATTCTGATGA TAATTAATGCC	Primer C, knock-out *tsiV1* in C6706
tsiV1-D	GGATCC ACACCTGCATCCTTAGCGCG	Primer D, knock-out *tsiV1* in C6706
tsiV1-E	GTGGAACATATGTTCATCGGGGG	Primer E, knock-out *tsiV1* in C6706
tsiV1-F	CCTTGAAGGAAAATCAGCAAGCC	Primer F, knock-out *tsiV1* in C6706
tsiV2-A	GGATC CAACCGATCTTGAAC	Primer A, knock-out *tsiV2* in C6706
tsiV2-B	TGAGCTATTCCTCTTTTAATTTATCAAT TAACATTTAA	Primer B, knock-out *tsiV2* in C6706
tsiV2-C	TTAAATGTTAATTGATAAATTAAAAGA GGAATAGCTCA	Primer C, knock-out *tsiV2* in C6706
tsiV2-D	GGATCC CTTATCTACTCGTTA	Primer D, knock-out *tsiV2* in C6706
tsiV3-A	GGATCC AGCATTGGCGCTGTT	Primer A, knock-out *tsiV3* in C6706
tsiV3-B	AATCCTAACTATTATCAACAAGCAAGT TATTCATTTTA	Primer B, knock-out *tsiV3* in C6706
tsiV3-C	TAAAATGAATAACTTGCTTGTTGATAA TAGTTAGGATT	Primer C, knock-out *tsiV3* in C6706
tsiV3-D	GGATCC AGCGCGAGATCAATAC	Primer D, knock-out *tsiV3* in C6706
tsiV3-E	GCAGGAGCTGATGGTGACTT	Primer E, knock-out *tsiV3* in C6706
tsiV3-F	CCTGTGCCACGATGACACTT	Primer F, knock-out *tsiV3* in C6706

vgrG-3-A	GGATCC CCACAAGTGAGCGTGCG	Primer A, knock-out *vgrG-3* in V52
vgrG-3-B	TTATTCATTTTATATCAACCTGTAACCT TGCCATGCTG	Primer B, knock-out *vgrG-3* in V52
vgrG-3-C	CAGCATGGCAAGGTTACAGGTTGATAT AAAATGAATAA	Primer C, knock-out *vgrG-3* in V52
vgrG-3-D	GGATCC GTAATGAAGATTTGATGAGG	Primer D, knock-out *vgrG-3* in V52
vgrG-3-E	GTT GCA CTG GGA ATG GTG GG	Primer E, knock-out *vgrG-3* in V52
vgrG-3-F	GGTCGAACCGTGATCATTTCGC	Primer F, knock-out *vgrG-3* in V52
tseL-A	GGATCC CGTTTAAACAGGCGGTGGCG	Primer A, knock-out *tseL* in V52
tseL-B	GATAACCATGATTTCACAGCAAACCTT ACC	Primer B, knock-out *tseL* in V52
tseL-C	GCTGTGAAATCATGGTTATCCCCTTAGT TC	Primer C, knock-out *tseL* in V52
tseL-D	GGATCC CACCGGCATTAATTATCATCAGATACC	Primer D, knock-out *tseL* in V52
tseL-E	GGCAACGCATTATTGCTTAGTGG	Primer E, knock-out *tseL* in V52
tseL-F	CTCGCCAAGTCCATGGTTGC	Primer F, knock-out *tseL* in V52
Ko1391-A	GCGGCCGC ACA ACA AGC CAC AGT GGG	Primer A, knock-out construct *VCV52_1391*
Ko1391-B	ATGGTTATCCCCTTAGTTCATAATGCGT TGCCATTCTT	Primer B, knock-out construct *VCV52_1391*
Ko1391-C	AAGAATGGCAACGCATTATGAACTAAG GGGATAACCAT	Primer C, knock-out construct *VCV52_1391*
Ko1391-D	GCGGCCGC CTACCGACATCATCCTCAGC	Primer D, knock-out construct *VCV52_1391*
ko1391-E	ATG TGA TTG CTG CGG GC	Primer E to confirm knock-out *VCV52_1391*
Ko1391-F	CCTCTCACACCTATAATCGCC	Primer F to confirm knock-out *VCV52_1391*

Ko1390-A	GGATCC AAGCGTGTTCACATTTGGAGCC	Primer A, knock-out construct *VCV52_1390*
Ko1390-B	CATTCTTGAGGATTATGCTCCGCTAATG TCGCCATCCTG	Primer B, knock-out construct *VCV52_1390*
Ko1390-C	CAGGATGGCGACATTAGCG GAGCATAATCCTCAAGAATG	Primer C, knock-out construct *VCV52_1390*
Ko1390-D	GCGGCCGC AGAAGCATTGACATCTATCCC	Primer D, knock-out construct *VCV52_1390*
Ko1390-E	GTTCGGTTGTCGTGCCCG	Primer E to confirm knock-out *VCV52_1390*
Ko1390-F	CGACAACAACGCAAAGGTATAACTATT CACC	Primer F to confirm knock-out *VCV52_1390*
1391 F	CCATGGCAACGCATTATTGCTT	To clone HIS-tagged *VCV52_1391*
1391 R	TCTAGATTAGTGGTGATGGTGATGATG TCCCCTTAGTTCAAATTGACG	To clone HIS-tagged *VCV52_1391*
1501 F	GAATTCACCATGATTCAGCAGTGGCTA ATGG	Forward primer for HIS-tagged *A55_1501*
1501 R	GGTACCTTAGTGGTGATGGTGATGATG TGCTAATTTCTCCCGAATTTGC	Reverse primer for HIS-tagged *A55_1501*
1345* F	GAATTCACCATGGCAACGCATTATTGC TT	Forward primer for HIS-tagged *A33_1345**
1501 ko A	CCCGGG TTCTTTCTTAAACGGTGACC	Primer A, knock-out construct *A55_1501*
1501 ko B	CCTGAAGTTCTTGTTGATTATCTAATAA CCAGAAGGTG	Primer B, knock-out construct *A55_1501*
1501 ko C	CACCTTCTGGTTATTAGATAATCAACA AGAACTTCAGG	Primer C, knock-out construct *A55_1501*
1501 ko D	ACTAGT TTACTTGCACTACCCGCAACCG	Primer D, knock-out construct *A55_1501*
1501 ko E	TGGGACCGCTACTCGAACGG	Primer E to confirm knock-out *A55_1501*
1501 ko F	TTACCAATTCCTAAGCCTTGCCCG	Primer F to confirm knock-out *A55_1501*
1501-3 ko A	ACTAGT TACAAGAGCACCTTGACCCCG	Primer A, knock-out construct *A55_1501-*

		1503
1501-3 ko B	AAGTTACTCGCCAAGCGTTGGTGGGTG TTTTTTGCTCC	Primer B, knock-out construct *A55_1501-1503*
1501-3 ko C	GGAGCAAAAAACACCCACCAACGCTTG GCGAGTAACTT	Primer C, knock-out construct *A55_1501-1503*
1501-3 ko D	CCCGGG ATTGGTTTGATGTTCTTTGCC	Primer D, knock-out construct *A55_1501-1503*
1501-3 ko E	CTACAAAGACGATGTGAACGGTGCC	Primer E to confirm knock-out *A55_1501-1503*
1501-3 ko F	TGACCACCTACGCCTTTGCC	Primer F to confirm knock-out *A55_1501-1503*
1503 fw	GAATTC ACC ATGAAAAGGTTGTTTCTAATACTATCC	Forward primer for *A55_1503*
1503 rev	TCTAGATTTGTCGTCGTCGTCTTTATAG TCCTCGCCAAGCGTTTTAATC	Reverse primer for FLAG-tagged *A55_1503*
1391 fw	A GAATTC ACC ATG GCAACGCATTATTGCTTAGTGGAAAAT GG	Forward primer for *VCV52_1391*
1391 rev	TTCTAGATTActtgtcatcgtcgtccttgtagtcTCCC CTTAGTTCAAATTGACGAGTTAAAGC	Reverse primer for FLAG-tagged *VCV52_1391*
vgrG-1 fw	ACCATGGCGACATTAGCGTACAGCATT GAGGTGGAAGGGCTGGAAGATGAGAC TCTGGTGGTCAGAGG	Forward primer to create truncated versions of *vgrG-1* (VCV52_1390)
vgrG-1 rev1	ATCTAGACTAAGCAATAATGCGTTGCC ATTCTTGAGGATTATGCTCTTTAACCC	Reverse primer to amplify full-length *vgrG-1* (VCV52_1390)
vgrG-1 rev2	ATCTAGACTAGGCTAAAGGACACACCT TCACCGCAGGAACATTCGCTTGCTCGG CC	Reverse primer to amplify basepairs 1-2037 of *vgrG-1* (VCV52_1390)
vgrG-1	ATCTAGACTACGGTAAGGCGGGCATTG	Reverse primer to

rev3	CACCGCCAAAGC	amplify basepairs 1-1926 of *vgrG-1* (VCV52_1390)
1391 fw1	A GAATTC ACC ATG gac tac aag gac gac gat gac aag GCAACGCATTATTGCTTAGTGGAAAATGG	Forward primer for full length *VCV52_1391*
1391 fw2	A GAATTC ACC ATG gac tac aag gac gac gat gac aag CTTTTACAATGGTTGAGTGAACAGACATCAAACG	Forward primer for *VCV52_1391* starting at basepair 43
1391 bw1	GGTCTAGATTATCCCCTTAGTTCAAATTGACGAGTTAAAGC	Reverse primer for *VCV52_1391*
ACD-A	AGC GGC CGC GCT ACT CGA ACG G	ACD knock-in into 1587
ACD-B	AGGCGTTGCTTCTTGAGCGGCTAAAGGACACACCTTCA	ACD knock-in into 1587
ACD-C	TGAAGGTGTGTCCTTTAGC CGCTCAAGAAGCAACGCCTGCGGTGAATAGTATCGCG	ACD knock-in into 1587
ACD-D	TTAGCCACTGCTGAATCATG CAATAATGCGTTGCC	ACD knock-in into 1587
ACD-E	AATGGCAACGCATTATTGCATGATTCAGCAGTGGCTAA	ACD knock-in into 1587
ACD-F	ACTAGT TTACTTGCACTACCCGCAACCG	ACD knock-in into 1587
ACD-G	ATG ATC GCC ACG GTC GTC G	ACD knock-in into 1587
ACD-H	TTACCAATTCCTAAGCCTTGCCCG	ACD knock-in into 1587

Table of accession numbers

(Strains for which sequences of T6SS gene clusters were generated in this study and are now available under the indicated accession codes are marked with an asterisk).

V. cholerae strain	GenBank Accession
12129(1)	ACFQ00000000
1587	AAUR01000000
2740-80	AAUT01000000
623-39	AAWG00000000
2010EL-1786	CP003069/CP003070
AM-19226	AATY01000000
Amazonia	AFSV00000000
BX 330286	ACIA00000000
C6706	AHGQ00000000
CA401*	KF228943/KF228947/KF228950
CIRS101	ACVW00000000
CP1041	SRA037374
DL4211*	KC955251/KF228941/KF228945
DL4215*	KF228942/KF228946/KF228949
HC-07A1	SRA035959
HC-32A1	SRA035998
HC-33A2	SRA035995
HC-38A1	SRA030739
HC-43B1	ALDP01000000
HE-25	ALEC00000000
HE-39	SRA030720
HE-45	ALED00000000
LMA3984-4	CP002555/CP002556
M66-2	NC_012578/NC_012580
MAK 757	AAUS00000000
MJ-1236	CP001485/CP001486

MO10	AAKF03000000
MZO-2	AAWF01000000
MZO-3	AAUU01000000
N16961	AE003852/AE003853
NIH41*	KF228944/KF228948/KF228951
O395	CP000626/CP000627
TM 11079-80	ACHW00000000
TMA21	ACHY00000000
V51	AAKI02000000
V52	AAKJ02000000
VL426	ACHV00000000

Table with locus tags and sequences for the bioinformatics analysis of Tap-1

Locus tags	*V. cholerae* strain	Nucleotide sequence (if necessary)
VC1417	N16961	-
VCV52_1391	V52	-
A33_1346*	AM-19226	ATGGCAACGCATTATTGCTTAGTGGAAAATGGAATGGAAAA GATGCTTTTACAATGGTTGAGTGAACAGACATCAAACGTCTA TTGGCTTGCTGATCATTCAACGTTTAAACAGGCGGTGGCGAA TAACAGTGGAATTGTATTTGATGGTACACAAGTTGTTTTCCAC GGAGAGACCTTTGCCCCTGTGATGGCGTTATCGCCTTGGTTAG TCCCTGTCTCTGATATGGTATCGGATATCGACTACGAGTGCTT ACAACAAGGTATATTCCTAAGTTGCTCCTGTCCCTCCACAGAG TTATTAAGCCATCTCCAAAGCCTTTTGATTGCAGCACTGGAAG GCGAAGAAGTACTGTTTCGTTTTTATGATAGGCAGGTCATTGT GCCTATGTTGGATGCGATGCGAGATCTGGAACGTAACGATTTT TTAGGGCCGGTAGAAAAATTAGCCGCCGTTAAGCAAGGAGTA TTTCAGGAATGGGGAAATACGCGTAGTTTCGAGTTTATCTATC AACCTGCCCCTTGGTGGAAAATTCAGCCTTATCATTTGATGCC ACTTTATCGAACCGAAGTGCATGCCCAAGTTTTAGAACGACGT TTTTGGGAAAAACTTCCTTACGTAATGGAACAGCTTGATGAGC CACAACAATGGATTAAAACCATTTTAGATGAAGCTAAGCAGG CTAATTTGGGGCACGATAATGCAGAGTACTTAGTTTTAAATCA CTTGTGGAAAGGCTCATTAACCACTCTAGAGCAGATGAGTGA TGATCTTCATCTTAATCAACAAGAACTTCAGGAAATAATGCA AATTCGGGAGAAATTAGCATGA
A55_1501	1587	-
A59_1447	623-39	-
VCG_002608*	12129(1)	ATGATTCAGCAGTGGCTAATGGAGCAAAAAACACCCACCTTC TGGTTATTAGATCAACAAGCTTTTAAATGCGTCATCGCAGCG AATAACGGAATAGGATTTGACGGCACTCAGATTCTATTTCAC GGCGAGACATTTGCTCCTGTGATGAGTTTGTCTCCTTGGTTAA TTCCTGTCTCCGACAAGGTGCTGGAATTGTCGGATGAGCTGTT ACAGCAAGGTATTTTCTTAACCAGTCATAGTTCTACAAGTGAA TTATTAAGCCATCTCCAAAGCCTTTTGGTTGCAGCACTGGAAG GCGAAGAAGTACTGTTTCGTTTTTATGATAGGCAGGTCATTTT GCCTATGCTGGATGCGATGCGAGATCTGGAACGTAACGATTTT TTAGGGCCGGTAGAAAAATTAGCCGCCGTTAAGCAAGGAGTAT TTCAGGAATGGGGAAATACGCGTAGTTTCGAGTTTATCTATCAA CCTGCCCCTTGGTGGAAAATTCAGCCTTATCATTTGATGCCACTT TATCGAACCGAAGTGCATGCCCAAGTTTTAGAACGACGTTTTTG GGAAAAGCTTCCTTACGTAATGGAACAGCTTGATGAGCCACAAC AATCGATTAAAACCATTTTAGATGAAGCTAAGAAGGCTAATTTG GGGCACGATAATGCAGAGTACTTAGTTTTAAATCACTTGTGGA AAGCCTCATTAACCACTCTAGAGCAGATGAGTGATGCTCTTCA TCTTAATCAACAAGAACTTCAGGAAATAATGCAAATTCGGGAG AAATTAGCATGA
VCV51_B0075	V51	-
No locus tag	Amazonia	ATGATTCAGCAGTGGCTAATGGAGCAAAAAATACCCACC TTCTGGTTATTAGATCAACAAGCTTTTAAATGCGTCATCGC AGCGAATAACGGAATAGGATTTGACGGCACTCAGATTCTA TTTCACGGCGAGACATTTGCTCCTGTGATGAGTTTGTCTCC TTGGTTAATTCCTGTCTCCGACAAGGTGCTGGAATTGTCGG ATGAGCTGTTACAGCAAGGTATTTTCTTAACCAGTCATAGTT CTACAAGTGAATTATTAAGCCATCTCCAAAGCCTTTTGGTTG CAGCACTGGAAGGCGAAGAAGTACTGTTTCGTTTTTATGATA

		GGCAGGTCATTTTGCCTATGCTGGATGCGATGCGAGATCTGG AACGTAACGATTTTTTAGGGCCGGTAGAAAAATTAGCCGCCG TTAAGCAAGGAGTATTTCAGGAATGGGGAAATACGCGTAGTT TCGAGTTTATCTATCAACCTGCCCCTTGGTGGAAAATTCAGCC TTATCATTTGATGCCACTTTATCGAACCGAAGTGCATGCCCAA GTTTTAGAACGACGTTTTTGGGAAAAGCTTCCTTACGTAATGG AACAGCTTGATGAGCCACAACAATGGATTAAAACCATTTTAG ATGAAGCTAAGCAGGCTAATTTGGGGCACGATAATGCAGAGT ACTTAGTTTTAAATCACTTGTGGAAAGCCTCATTAACCACTCT AGAGCAGATGAGTGATGCTCTTCATCTTAATCAACAAGAACTT CAGGAAATAATGCAAATTCGGGAGAAATTAGCATGA
VIF_002053	TM11079-80	-
VCHC43B1_1452	HC-43B1	-
VCHE45_1517	HE-45	-
VCHE39_2321	HE-39	-
A51_B1431	MZO-3	-
VCB_003440	TMA21	-
No locus tag	DL4215	ATGGCAACGCATTATTGCTTAGTGGAAAATGGAATGGAAAAGA TGCTTTTACAATGGTTGAGTGAACAGACATCAAACGTCTATTGG CTTGCTGATCATTCAACGTTTAAACAGGCGGTGGCGAATAACAG TGGAATTGTATTTGATGGTACACAAGTTGTTTTCCACGGCGAGAC CTTTGCTCCTGTGATGAGTTTGTCCCCTTGGCTAATTCCTGTCTCA GATAAGGTGTTGGCATTACCGGATGAGCTGTTACAGCAAGGCAT TTTCTTAACCAGTCATAGTTCTACAAGTGAATTATTAAGCCATCT TCAAAGCCTTTTGATTGCAGCACTGGAAGGCGAAGAAGTACTGT TTCGTTTTTATGATAGGCAGGTCATTGTGCCTATGCTGGATGCGA TGCGAGATCTGGAACGTAACGATTTTTTAGGGCCGGTAGAAAAA TTAGCCGCCGTTAAGCAAGGAGTATTTCAGGAATGGGGAAATAC GCGTAGTTTCGAGTTTATCTATCAACCTGCCCCTTGGTGGAAAAT TCAGCCTTATCATTTGATGCCACTTTATCGAACCGAAGTGCATGC CCAAGTTTTAGAACGACGTTTTTGGGAAAAGCTTCCTTACGTAAT GGAACAGCTTGATGAGCCACAACAATCGATTAAAACCATTTTAG ATGAAGCTAAGCAGGCTAATTTGGGGCACGATAATGCAGAGTA CTTAGTTTTAAATCACTTGTGGAAAGCCTCATTAGCCACTCTAG AGCAGATGAGTGATGCTCTTCATCTTAATCAACAAGAACTTCA GGAAATAATGCAAATTCGGGAGAAATTAGCATGA
VCA_001063	VL426	-
No locus tag	DL4211	ATGGCAACGCATTATTGCTTAGTGGAAAATGGAATGGAAA AGATGCTTTTACAATGGTTGAGTGAACAGACATCAAACGTC TATTGGCTTGCTGATCATTCAACGTTTAAACAGGCGGTGGCG AATAACAGCGGAATTGTATTTGATGGTACACAAGTTGTTTTC CACGGAGAGACCTTTGCCCCTGTGATGGCGTTATCGCCTTGGT TAGTCCCTGTCTCTGATATGGTATCGGATATCGACTACGAGTG CTTATTACAACAAGGTATATTCCTAAGTTGCTCCTGTCCCTCC ACAGAGTTATTAAGCCATCTCCAAAGCCTTTTGATTGCAGCAC TGGAAGGCGAAGAAGTACTGTTTCGTTTTTATGATAGGCAGGT CATTGTGCCTATGTTGGATGCGATGCGAGATCTGGAACGCAA TGATTTTTTAGGGCCGGTAGAGAAATTAGCCGCCGTTAAGCA AGGAGTATTTCAGGAATGGGGAAATACGCGCAGTTCCGAGTT TATCTATCAGCCAGCTCCTTGGTGGAAAATACAACCTTATCAT TTGATGCCACTTTATCGAACTGAAGTACATGCCCAAGTTTTAG AACGACGTTTTTGGGAAAAGCTTCCTTACGTAATGGAACAGCT TGATGAGCCACATCAATGGATTAAAACCATTTTAGATGAAGCT AAGCAGGCTAATTTGGGGCACGATAATGCAGAGTACTTAGTT TTAAATCACTTGTGGAAAGGCTCATTAACCACTCTAGAGCAG ATGAGTGATGCTCTTCATCTTAATCAACAAGAACTTCAGGAA ATAATGCAAATTCGGGAGAAATTAGCATGA
No locus tag	HE-25	ATGGCAACGCATTATTGCTTAGTGGAAAATGGAATGGAAAAGA

		TGCTTTTACAATGGTTGAGTGAACAGACATCAAACGTCTATTGG CTTGCTGATCATTCAACGTTTAAACAGGCGGTGGCGAATAACAG TGGAATTGTATTTGATGGTACACAAGTTGTTTTCCACGGAGAGA CCTTTGCCCCTGTGATGGCGTTATCGCCTTGGTTAGTCCCTGTCT CTGATATGGTATCGGATATCGACTACGAGTGCTTACAACAAGGT ATATTCCTAAGTTGCTCCTGTCCCTCCACAGAGTTATTAAGCCAT CTCCAAAGCCTTTTGATTGCAGCACTGGAAGGCGAAGAAGTACT GTTTCGTTTTTATGATAGGCAGGTCATTGTGCCTATGTTGGATGC GATGCGAGATCTGGAACGTAACGATTTTTTAGGGCCGGTAGAAA AATTAGCCGCCGTTAAGCAAGGAGTATTTCAGGAATGGGGAAAT ACGCGTAGTTTCGAGTTTATCTATCAACCTGCCCCTTGGTGGAAA ATTCAGCCTTATCATTTGATGCCACTTTATCGAACCGAAGTGCAT GCCCAAGTTTTAGAACGACGTTTTTGGGAAAAACTTCCTTACGT AATGGAACAGCTTGATGAGCCACAACAATGGATTAAAACCATTT TAGATGAAGCTAAGCAGGCTAATTTGGGGCACGATAATGCAGA GTACTTAGTTTTAAATCACTTGTGGAAAGGCTCATTAACCACT CTAGAGCAGATGAGTGATGATCTTCATCTTAATCAACAAGAA CTTCAGGAAATAATGCAAATTCGGGAGAAATTAGCATGA
Vch1786_I0916	2010EL-1786	-
No locus tag	2740-80	ATGGCAACGCATTATTGCTTAGTGGAAAATGGAATGGAAAAG ATGCTTTTACAATGGTTGAGTGAACAGACATCAAACGTCTATT GGCTTGCTGATCATTTAACGTTTAAACAGGCGGTGGCGAATAA CAGTGGAATTGTATTTGATGGTACACAAGTTGTTTTCCACGGAG AGACCTTTGCCCCTGTGATGGCGTTATCGCCCTGGTTAGTCCCTG TCTCTGATATGGTATCGGATATCGACTACGAGTGCTTACAACAA GGTATATTCCTAAGTTGCTCCTGTCCCTCCACAGAGTTATTAAGC CATCTCCAAAGCCTTTTGATTGCAGCACTGGAAGGCGAAGAAGT ACTGTTTCGTTTTTATGATAGGCAGGTCATTGTGCCTATGCTTGA ACGAATGGATGAAATAGAAAAAAATCAGTTCTTAGGTAAGATA AATCAGTTAGTCGTGTTCGATAGCCATGAGAAAAATCGGTTAGT CTCATTTACCAATACATCAGACGCTCAGTTTACGGCCAAAAAGA CAACGTGGTGGGTGATCAAACGACACCATTTGGAGGCTCATGAG AATTTGCCTCTTGTTCAAGCTAACCTCGAATCTTGGCTATGGCGG CATTTTCCAAAAGTGATGAACGAGTACTTAATTCAAGGGCGTGA TATCGCTTCCATGCTCGCTCCATCTCTCTCTGTTCCTGAACAAAC CTTAACATATCGTGTGTTAAGTGCCGCTATCGTTGCTGTTGTCG GTGCCGAACAATTGACTCAGCCTAGCATGATTGAATTTCTACG AGACTATAACAACGAAGAAGCTCAGCTCGCTCTGCATGCTTTA ACTCGTCAATTTGAACTAAGGGGATAA
VCF_000214	BX 330286	-
No locus tag	C6706	ATGGCAACGCATTATTGCTTAGTGGAAAATGGAATGGAAAAGA TGCTTTTACAATGGTTGAGTGAACAGACATCAAACGTCTATTGGC TTGCTGATCATTTAACGTTTAAACAGGCGGTGGCGAATAACAGTG GAATTGTATTTGATGGTACACAAGTTGTTTTCCACGGAGAGACCT TTGCCCCTGTGATGGCGTTATCGCCCTGGTTAGTCCCTGTCTCTGA TATGGTATCGGATATCGACTACGAGTGCTTACAACAAGGTATAT TCCTAAGTTGCTCCTGTCCCTCCACAGAGTTATTAAGCCATCTCCA AAGCCTTTTGATTGCAGCACTGGAAGGCGAAGAAGTACTGTTTCG TTTTTATGATAGGCAGGTCATTGTGCCTATGCTTGAACGAATGGA TGAAATAGAAAAAAATCAGTTCTTAGGTAAGATAAATCAGTTAG TCGTGTTCGATAGCCATGAGAAAAATCGGTTAGTCTCATTTACCA ATACATCAGACGCTCAGTTTACGGCCAAAAAGACAACGTGGTGG GTGATCAAACGACACCATTTGGAGGCTCATGAGAATTTGCCTCTT GTTCAAGCTAACCTCGAATCTTGGCTATGGCGGCATTTTCCAAAA GTGATGAACGAGTACTTAATTCAAGGGCGTGATATCGCTTCCATG CTCGCTCCATCTCTCTCTGTTCCTGAACAAACCTTAACATATCGT GTGTTAAGTGCCGCTATCGTTGCTGTTGTCGGTGCCGAACAAT TGACTCAGCCTAGCATGATTGAATTTCTACGAGACTATAACAA

		CGAAGAAGCTCAGCTCGCTCTGCATGCTTTAACTCGTCAATTTG AACTAAGGGGATAA
No locus tag	CA401	ATGGCAACGCATTATTGCTTAGTGGAAAATGGAATGGAAAAG ATGCTTTTACAATGGTTGAGTGAACAGACATCAAACGTCTATT GGCTTGCTGATCATTTAACGTTTAAACAGGCGGTGGCGAATAA CAGTGGAATTGTATTTGATGGTACACAAGTTGTTTTCCACGGAG AGACCTTTGCCCCTGTGATGGCGTTATCGCCCTGGTTAGTCCCTG TCTCTGATATGGTATCGGATATCGACTACGAGTGCTTACAACAA GGTATATTCCTAAGTTGCTCCTGTCCCTCCACAGAGTTATTAAGC CATCTCCAAAGCCTTTTGATTGCAGCACTGGAAGGCGAAGAAGT ACTGTTTCGTTTTTATGATAGGCAGGTCATTGTGCCTATGCTTGA ACGAATGGATGAAATAGAAAAAAATCAGTTCTTAGGTAAGATAA ATCAGTTAGTCGTGTTCGATAGCCATGAGAAAAATCGGTTAGT CTCATTTACCAATACATCAGACGCTCAGTTTACGGCCAAAAAG ACAACGTGGTGGGTGATCAAACGACACCATTTGGAGGCTCATG AGAATTTGCCTCTTGTTCAAGCTAACCTCGAATCTTGGCTATGG CGGCATTTTCCAAAAGTGATGAACGAGTACTTAATTCAAGGGCG TGATATCGCTTCCATGCTCGCTCCATCTCTCTCTGTTCCTGAACA AACCTTAACATATCGTGTGTTAAGTGCCGCTATCGTTGCTGTTG TCGGTGCCGAACAATTGACTCAGCCTAGCATGATTGAATTTCT ACGAGACTATAACAACGAAGAAGCTCAGCTCGCTCTGCATGC TTTAACTCGTCAATTTGAACTAAGGGGATAA
VCH002160*	CIRS101	ATGGCAACGCATTATTGCTTAGTGGAAAATGGAATGGAAAAGAT GCTTTTACAATGGTTGAGTGAACAGACATCAAACGTCTATTGGCT TGCTGATCATTTAACGTTTAAACAGGCGGTGGCGAATAACAGTGG AATTGTATTTGATGGTACACAAGTTGTTTTCCACGGAGAGACCTTT GCCCCTGTGATGGCGTTATCGCCCTGGTTAGTCCCTGTCTCTGATA TGGTATCGGATATCGACTACGAGTGCTTACAACAAGGTATATTCC TAAGTTGCTCCTGTCCCTCCACAGAGTTATTAAGCCATCTCCAAAG CCTTTTGATTGCAGCACTGGAAGGCGAAGAAGTACTGTTTCGTTTT TATGATAGGCAGGTCATTGTGCCTATGCTTGAACGAATGGATGAA ATAGAAAAAAATCAGTTCTTAGGTAAGATAAATCAGTTAGTCGTG TTCGATAGCCATGAGAAAAATCGGTTAGTCTCATTTACCAATACA TCAGACGCTCAGTTTACGGCCAAAAAGACAACGTGGTGGGTGATC AAACGACACCATTTGGAGGCTCATGAGAATTTGCCTCTTGTTCAA GCTAACCTCGAATCTTGGCTATGGCGGCATTTTCCAAAAGTGATG AACGAGTACTTAATTCAAGGGCGTGATATCGCTTCCATGCTCGCT CCATCTCTCTCTGTTCCTGAACAAACCTTAACATATCGTGTGTTA AGTGCCGCTATCGTTGCTGTTGTCGGTGCCGAACAATTGACTCA GCCTAGCATGATTGAATTTCTACGAGACTATAACAACGAAGAA GCTCAGCTCGCTCTGCATGCTTTAACTCGTCAATTTGAACTAAG GGGATAA
VCHC7A1_02526	HC-07A1	-
No locus tag	HC-32A1	ATGGCAACGCATTATTGCTTAGTGGAAAATGGAATGGAAAAGATG CTTTTACAATGGTTGAGTGAACAGACATCAAACGTCTATTGGCTT GCTGATCATTTAACGTTTAAACAGGCGGTGGCGAATAACAGTGG AATTGTATTTGATGGTACACAAGTTGTTTTCCACGGAGAGACCTTT GCCCCTGTGATGGCGTTATCGCCCTGGTTAGTCCCTGTCTCTGATA TGGTATCGGATATCGACTACGAGTGCTTACAACAAGGTATATTCC TAAGTTGCTCCTGTCCCTCCACAGAGTTATTAAGCCATCTCCAAA GCCTTTTGATTGCAGCACTGGAAGGCGAAGAAGTACTGTTTCGTT TTTATGATAGGCAGGTCATTGTGCCTATGCTTGAACGAATGGATG AAATAGAAAAAAATCAGTTCTTAGGTAAGATAAATCAGTTAGTC GTGTTCGATAGCCATGAGAAAAATCGGTTAGTCTCATTTACCAAT ACATCAGACGCTCAGTTTACGGCCAAAAAGACAACGTGGTGGGT GATCAAACGACACCATTTGGAGGCTCATGAGAATTTGCCTCTTGT TCAAGCTAACCTCGAATCTTGGCTATGGCGGCATTTTCCAAAAGT GATGAACGAGTACTTAATTCAAGGGCGTGATATCGCTTCCATGCT CGCTCCATCTCTCTCTGTTCCTGAACAAACCTTAACATATCGTGT GTTAAGTGCCGCTATCGTTGCTGTTGTCGGTGCCGAACAATTGA CTCAGCCTAGCATGATTGAATTTCTACGAGACTATAACAACGA

		AGAAGCTCAGCTCGCTCTGCATGCTTTAACTCGTCAATTTGAAC TAAGGGGATAA
No locus tag	HC-33A2	ATGGCAACGCATTATTGCTTAGTGGAAAATGGAATGGAAAAGA TGCTTTTACAATGGTTGAGTGAACAGACATCAAACGTCTATTGG CTTGCTGATCATTTAACGTTTAAACAGGCGGTGGCGAATAACAG TGGAATTGTATTTGATGGTACACAAGTTGTTTTCCACGGAGAGA CCTTTGCCCCTGTGATGGCGTTATCGCCCTGGTTAGTCCCTGTCT CTGATATGGTATCGGATATCGACTACGAGTGCTTACAACAAGGT ATATTCCTAAGTTGCTCCTGTCCCTCCACAGAGTTATTAAGCCAT CTCCAAAGCCTTTTGATTGCAGCACTGGAAGGCGAAGAAGTACT GTTTCGTTTTTATGATAGGCAGGTCATTGTGCCTATGCTTGAACG AATGGATGAAATAGAAAAAAATCAGTTCTTAGGTAAGATAAATC AGTTAGTCGTGTTCGATAGCCATGAGAAAAATCGGTTAGTCTCAT TTACCAATACATCAGACGCTCAGTTTACGGCCAAAAAGACAACG TGGTGGGTGATCAAACGACACCATTTGGAGGCTCATGAGAATTT GCCTCTTGTTCAAGCTAACCTCGAATCTTGGCTATGGCGGCATTTT CCAAAAGTGATGAACGAGTACTTAATTCAAGGGCGTGATATCGC TTCCATGCTCGCTCCATCTCTCTCTGTTCCTGAACAAACCTTAACA TATCGTGTGTTAAGTGCCGCTATCGTTGCTGTTGTCGGTGCCGA ACAATTGACTCAGCCTAGCATGATTGAATTTCTACGAGACTA TAACAACGAAGAAGCTCAGCTCGCTCTGCATGCTTTAACTCG TCAATTTGAACTAAGGGGATAA
VCHC38A1_1503	HC-38A1	-
VCM66_1372	M66-2	-
A53_01538	MAK757	-
VCD002930	MJ-1236	-
A5A_1699	MZO-2	-
VC395_1536	O395	-
VCCP104114_2232	CP1041	-
VchoM_00682	MO10	-

List with locus tags and sequences for the analysis of the effector modules

Locus tags of the effector-immunity pairs are shown in the same order than the strain names are listed in figure S1(Unterweger et al., 2014):

Auxiliary cluster 1

Strain name	(locus tag effector-encoding gene/ immunity protein-encoding gene)
V52	(VCV52_1392/VCV52_1393)
2010EL-1789	(Vch1786_I0917/Vch1786_I0918)
2740-80	(VC274080_1493/VC274080_1494)
BX330286	(VCF_000213/VCF_000212)
C6706	(no locus tag available/no locus tag available)
CIRS101	(VCH_002161/VCH_002162)
CP1041	(VCCP104114_2233/VCCP104114_2234)
HC-07A1	(VCHC7A1_02527/VCHC7A1_02528)
HC-32A1	(VCHC32A1_1498/VCHC32A1_1499)
HC-33A2	(VCHC33A2_1482/VCHC33A2_1483)
HC-38A1	(VCHC38A1_1504/VCHC38A1_1505)
M66-2	(VCM66_1373/VCM66_1374)
MAK757	(A53_01539/A53_01540)
MJ-1236	(VCD_002929/VCD_002928)
MO-10	(VchoM_00683/VchoM_00684)
N16961	(VC1418/VC1419)
NIH41	(no locus tag available/protein_id="AHX36912.1")
CA401	(protein_id="AHX36907.1"/protein_id="AHX36908.1")
O395	(VC395_1537/VC395_1538)
MZO-2	(A5A_1700/A5A_1701)
LMA3984-4	(VCLMA_A1247/VCLMA_A1248)
AM19226	(A33_1346/A33_1347)
MZO-3	(A51_B1432/A51_B1433)
12129(1)	(VCG_002607/VCG_002606)
HE-25	(VCHE25_2421/VCHE25_2422)
TMA21	(VCB_003441/VCB_003442)
1587	(A55_1502/A55_1503)
DL4215	(protein_id="AHX36901.1"/protein_id="AHX36902.1")
TM11079-80	(VIF_002054/VIF_002055)
V51	(VCV51_B0076/ VCV51_B0077)
623-39	(A59_1448/A59_1449)
Amazonia	(no locus tag available/no locus tag available)
DL4211	(protein_id="AHM24037.1"/protein_id="AHM24038.1")
HE-45	(VCHE45_1518/VCHE45_1519)
HC-43B1	(VCHC43B1_1453/VCHC43B1_1454)
HE-39	(VCHE39_2323/VCHE39_2324)
VL426	(VCA_001062/VCA_001061)

Auxiliary cluster 2:

V52 (VCV52_A0043/VCV52_A0044)
2010EL-1789 (Vch1786_II0816/Vch1786_II0817)
2740-80 (VC274080_A0056/VC274080_A0057)
BX330286 (VCF_002023/VCF_002022)
C6706 (no locus tag available/no locus tag available)
CIRS101 (VCH_002522/VCH_002521)
CP1041 (VCCP104114_1116/VCCP104114_1115)
HC-07A1 (VCHC7A1_01128/VCHC7A1_01129)
HC-32A1 (VCHC32A1_3612/ VCHC32A1_3613)
HC-33A2 (VCHC33A2_3581/VCHC33A2_3582)
HC-38A1 (VCHC38A1_2808/VCHC38A1_2809)
M66-2 (VCM66_A0019/VCM66_A0020)
MAK757 (A53_03634/A53_03635)
MJ-1236 (VCD_000213/VCD_000212)
MO-10 (VchoM_02434/VchoM_02435)
N16961 (VCA0020/VCA0021)
MZO-2 (no locus tag available/no locus tag available)
NIH41 (protein_id="AHX36804.1"/protein_id="AHX36805.1")
CA401 (protein_id="AHX36793.1"/protein_id="AHX36794.1")
O395 (VC395_A0017/VC395_A0018)
12129(1) (VCG_001064/VCG_001063)
HC-43B1 (VCHC43B1_2096/VCHC43B1_2097)
LMA3984-4 (VCLMA_B0017/VCLMA_B0018)
VL426 (VCA_000458/VCA_000457)
TMA21 (VCB_002186/VCB_002187)
Amazonia (no locus tag available/no locus tag available)
TM11079-80 (VIF_003265/VIF_003266)
HE-45 (VCHE45_2803/VCHE45_2802)
HE-39 (VCHE39_0208/VCHE39_0209)
623-39 (A59_A0182/A59_A0183)
1587 (A55_A0021/A55_A0022)
V51 (VCV51_A0204/VCV51_A0205)
AM19226 (A33_A0310/A33_A0311)
HE-25 (VCHE25_0876/VCHE25_0877)
MZO-3 (A51_C0018/A51_C0019)
DL4215 (protein_id="AHX36786.1"/protein_id="AHX36787.1")
DL4211 (protein_id="AHX36780.1"/protein_id="AHX36781.1")

Large cluster:

V52 (VCV52_A0139/VCV52_A0140)
2010EL-1789 (Vch1786_II0916/Vch1786_II0917)
2740-80 (VC274080_A0152/VC274080_A0153)

BX330286 (VCF_001931/not annotated)
C6706 (no locus tag available/no locus tag available)
CIRS101 (VCH_002428/not annotated)
CP1041 (VCCP104114_1022/not annotated)
HC-07A1 (VCHC7A1_01223/VCHC7A1_01224)
HC-32A1 (VCHC32A1_3705/VCHC32A1_3706)
HC-33A2 (VCHC33A2_3675/not annotated)
HC-38A1 (VCHC38A1_2902/VCHC38A1_2903)
M66-2 (VCM66_A0121/VCM66_A0122)
MAK757 (A53_03730/A53_03731)
MJ-1236 (VCD_000122/not annotated)
MO-10 (VchoM_02536/VchoM_02537)
N16961 (VCA0123/VCA0124)
NIH41 (protein_id="AHX36893.1"/protein_id="AHX36894.1")
CA401 (protein_id="AHX36872.1"/protein_id="AHX36873.1")
O395 (VC0395_0014/VC0395_0013)
VL426 (VCA_000362/not annotated)
V51 (VCV51_A0278/not annotated)
MZO-3 (A51_C0113/A51_C0114)
HE-39 (VCV51_A0278/not annotated)
MZO-2 (A5A_A0193/A5A_A0194)
623-39 (A59_A0283/not annotated)
1587 (A55_A0115/not annotated)
12129(1) (VCG_000967/not annotated)
DL4215 (protein_id="AHX36850.1"/protein_id="AHX36851.1")
LMA3984-4 (VCLMA_B0110/not annotated)
HC-43B1 (VCHC43B1_2186/not annotated)
HE-45 (VCHE45_2712/not annotated)
DL4211 (protein_id="AHX36826.1"/protein_id="AHX36827.1")
AM19226 (A33_A0407/A33_A0408)
TMA21 (VCB_002278/not annotated)
HE-25 (VCHE25_0973/not annotated)
Amazonia (no locus tag available/no locus tag available)
TM11079-80 (VIF_003361/not annotated)

Sequences of the immunity proteins, which we used to group the effector modules into families:

Auxiliary cluster 1

>623-39

MKKTLVILSLFLVNACTSNAASVPGEPWYVGVTMPSFYPVRVTQVYGVNNKED
WTVLIHNFMTFVSDSDVEYIRNRFPDYDGFGLPHALGITSRNQIGTGTNHLPDTL
YVYWVSLFDTKFYVTKYDIPDNVKQLSATEVGYTRSDGEFFESCYRTRFVFGLL
PNGQAKVWLSGCGETIYLTELAPDEILDRDSNGVKFDSYRKSSSFADVQQRAKD
AGVELEPIPWDKVNKVYSINEIKTLGE

>1587

MKRLFLILSIVLLSACSSNAASVPGKAWYVGVTMPSFYPVKITQIYGVNEKENWT
VLIHNIIPRMSDSDIEYIRNRFPEYDGFGLPHALGITSRNQIGIGTNHLPDTLYVHW
VSLFDNKFYVTKYDIPDNVKQLSATEVSYTRSDGAVFESCYRTRFVFGLLPNGRA
KVWLSDCGETIYLTELAPDQTPDRDSNGFKADTYKESSYISNIQQRAKDAGVELE
PIPWDKVNKVYSINEIKTLGE

>2010EL-1786

MKLLNNLAIKKLISVVVLFIALFFLGYDYYQTSQPNIWGDEPDESYITISGKKPIDA
YLEVWTHFWVTGDECEAYSYDLFGQKAHQGGKISQKITHDFAKDGSNFEFRIPY
QTYKNSQNCIVELRDFSIQAHNDFDTVGFAQLRFSPAGREYYNREVDLNSLITAD
NCNSDIFKSIRKEWAGAIGCHFYVDGKKKTKDEEFNAYTIYYDFSKFNNDTVIHY
DILAGENYRSEPLSSAQKTAVVSDDN

>2740-80

MKLLNNLAIKKLISVVVLFIALFFLGYDYYQTSQPNIWGDEPDESYITISGKKPIDA
YLEVWTHFWVTGDECEAYSYDLFGQKAHQGGKISQKITHDFAKDGSNFEFRIPY
QTYKNSQNCIVELRDFSIQAHNDFDTVGFAQLRFSPAGREYYNREVDLNSLITAD
NCNSDIFKSIRKEWAGAIGCHFYVDGKKKTKDEEFNAYTIYYDFSKFNNDTVIHY
DILAGENYRSEPLSSAQKTAVVSDDN

>12129(1)

MKRLFLILSIVLLSACSSNAASVPGKAWYVGVTMPSFYPVKITQIYGVNEKENWT
VLIHNIIPRMSDSDIEYIRNRFPEYDGFGLPHALGITSRNQIGIGTNHLPDTLYVHW
VSLFDNKFYVTKYDIPDNVKQLSATEVSYTRSDGAVFESCYRTRFVFGLLPNGRA
KVWLSDCGETIYLTELAPDQTPDRDSNGFKADTYKESSYIRNIQQRAKDAGVELE
PIPWDKVNKVYSINEIKTLGE

>AM19226

MKRLFLILSIVLLSACSSNAASVPGKAWYVGVTMPSFYPVKITQIYGVNEKENWT
VLIHNIIPRMSDSDIEYIRNRFPEYDGFGLPHALGITSRNQIGIGTNHLPDTLYVHW
VSLFDNKFYVTKYDIPDNVKQLSATEVSYTRSDGAVFESCYRTRFVFGLLPNGRA
KVWLSDCGETIYLTELAPDQTPDRDSNGFKADTYKESSYIRNIQQRAKDAGVELE
PIPWDKVNKVYSINEIKTLGE

>Amazonia

MKKTLVILSLFLVNACTSNAASVPGEPWYVGVTMPSFYPVRVTQVYGVNNKED
WTVLIHNFMTFVSDSDVEYIRNRFPDYDGFGLPHALGITSRNQIGTGTNHLPDTL
YVYWVSLFDTKFYVTKYDIPDNVKQLSATEVGYTRSDGEFFESCYRTRFVFGLL
PNGQAKVWLSGCGETIYLTELAPDEILDRDSNGVKFDSYRKSSSFADVQQRAKD
AGVELEPIPWDKVNKVYSINEIKTLGE

>BX_330286

MKLLNNLAIKKLISVVVLFIALFFLGYDYYQTSQPNIWGDEPDESYITISGKKPIDA
YLEVWTHFWVTGDECEAYSYDLFGQKAHQGGKISQKITHDFAKDGSNFEFRIPY
QTYKNSQNCIVELRDFSIQAHNDFDTVGFAQLRFSPAGREYYNREVDLNSLITAD
NCNSDIFKSIRKEWAGAIGCHFYVDGKKKTKDEEFNAYTIYYDFSKFNNDTVIHY
DILAGENYRSEPLSSAQKTAVVSDDN

>C6706_

MKLLNNLAIKKLISVVVLFIALFFLGYDYYQTSQPNIWGDEPDESYITISGKKPIDA
YLEVWTHFWVTGDECEAYSYDLFGQKAHQGGKISQKITHDFAKDGSNFEFRIPY
QTYKNSQNCIVELRDFSIQAHNDFDTVGFAQLRFSPAGREYYNREVDLNSLITAD
NCNSDIFKSIRKEWAGAIGCHFYVDGKKKTKDEEFNAYTIYYDFSKFNNDTVIHY
DILAGENYRSEPLSSAQKTAVVSDDN

>CA401

MKLLNNLAIKKLISVVVLFIALFFLGYDYYQTSQPNIWGDEPDESYITISGKKPIDA
YLEVWTHFWVTGDECEAYSYDLFGQKAHQGGKISQKITHDFAKDGSNFEFRIPY
QTYKNSQNCIVELRDFSIQAHNDFDTVGFAQLRFSPAGREYYNREVDLNSLFTAD
NCNSDIFKSIRKEWAGAIGCHFYVDGKKKTKDEEFNAYTIYYDFSKFNNDTVIHY
DILAGENYRSEPLSSAQKTAVVSDDN

>CIRS101

MKLLNNLAIKKLISVVVLFIALFFLGYDYYQTSQPNIWGDEPDESYITISGKKPIDA
YLEVWTHFWVTGDECEAYSYDLFGQKAHQGGKISQKITHDFAKDGSNFEFRIPY
QTYKNSQNCIVELRDFSIQAHNDFDTVGFAQLRFSPAGREYYNREVDLNSLITAD
NCNSDIFKSIRKEWAGAIGCHFYVDGKKKTKDEEFNAYTIYYDFSKFNNDTVIHY
DILAGENYRSEPLSSAQKTAVVSDDN

>CP1041

MKLLNNLAIKKLISVVVLFIALFFLGYDYYQTSQPNIWGDEPDESYITISGKKPIDA
YLEVWTHFWVTGDECEAYSYDLFGQKAHQGGKISQKITHDFAKDGSNFEFRIPY
QTYKNSQNCIVELRDFSIQAHNDFDTVGFAQLRFSPAGREYYNREVDLNSLITAD
NCNSDIFKSIRKEWAGAIGCHFYVDGKKKTKDEEFNAYTIYYDFSKFNNDTVIHY
DILAGENYRSEPLSSAQKTAVVSDDN

>DL4211

MILECLRCGTKKQAAIKELAIRASKENKMKKLFFLLLVLLVGSCSTNAAYETEQP
WTWRVGVTMPSFYPVNVTRVYGVNNKEDWTVLIHNFMTFVSDSDISYIRQIFPD
YDGFGLPHALGITSRSQIGVGTTQLPDTLYMYWVSLFDTKFYVTKYEIPDNVKQL

VATKTTYTKWNGFIVDPCYRAEFIFGLLPNGQAKVWLDGCGETIYLTELAPDQIL
DRDSNGVKFDSYRKSSSFADVQQRAKDAGVELEPIPWDKVNKVYSRYEIKTLGE

>DL4215

MKRLFLILSIVLLSACSSNAASVPGKAWYVGVTMPSFYPVKITQIYGVNEKENWT
VLIHNIIPRMSDSDIEYIRNRFPKYDGFGLPHALGITSRNQIGIGTNHLPDTLYVHW
VSLFDNKFYVTKYDIPDNVKQLSATEVSYTRSDGAVFESCYRTRFVFGLLPNGRA
KVWLSDCGETIYLTELAPDQTPDRDSNGFKADTYKESSYIRNIQQRAKDAGVELE
PIPWDKVNKVYSINEIKTLGE

>HC-07A1

MKLLNNLAIKKLISVVVLFIALFFLGYDYYQTSQPNIWGDEPDESYITISGKKPIDA
YLEVWTHFWVTGDECEAYSYDLFGQKAHQGGKISQKITHDFAKDGSNFEFRIPY
QTYKNSQNCIVELRDFSIQAHNDFDTVGFAQLRFSPAGREYYNREVDLNSLITAD
NCNSDIFKSIRKEWAGAIGCHFYVDGKKKTKDEEFNAYTIYYDFSKFNNDTVIHY
DILAGENYRSEPLSSAQKTAVVSDDN

>HC-32A1

MKLLNNLAIKKLISVVVLFIALFFLGYDYYQTSQPNIWGDEPDESYITISGKKPIDA
YLEVWTHFWVTGDECEAYSYDLFGQKAHQGGKISQKITHDFAKDGSNFEFRIPY
QTYKNSQNCIVELRDFSIQAHNDFDTVGFAQLRFSPAGREYYNREVDLNSLITAD
NCNSDIFKSIRKEWAGAIGCHFYVDGKKKTKDEEFNAYTIYYDFSKFNNDTVIHY
DILAGENYRSEPLSSAQKTAVVSDDN

>HC-33A2

MKLLNNLAIKKLISVVVLFIALFFLGYDYYQTSQPNIWGDEPDESYITISGKKPIDA
YLEVWTHFWVTGDECEAYSYDLFGQKAHQGGKISQKITHDFAKDGSNFEFRIPY
QTYKNSQNCIVELRDFSIQAHNDFDTVGFAQLRFSPAGREYYNREVDLNSLITAD
NCNSDIFKSIRKEWAGAIGCHFYVDGKKKTKDEEFNAYTIYYDFSKFNNDTVIHY
DILAGENYRSEPLSSAQKTAVVSDDN

>HC-38A1

MKLLNNLAIKKLISVVVLFIALFFLGYDYYQTSQPNIWGDEPDESYITISGKKPIDA
YLEVWTHFWVTGDECEAYSYDLFGQKAHQGGKISQKITHDFAKDGSNFEFRIPY
QTYKNSQNCIVELRDFSIQAHNDFDTVGFAQLRFSPAGREYYNREVDLNSLITAD
NCNSDIFKSIRKEWAGAIGCHFYVDGKKKTKDEEFNAYTIYYDFSKFNNDTVIHY
DILAGENYRSEPLSSAQKTAVVSDDN

>HC-43B1

MILECRRCGTKKQAAIKELAISASKENKMKKLFFLLLVLLVGSCSTNAAYETEQP
WTWRVGVTMPSFYPVNVTRVYGVNNKEDWTVLIHNFMTFVSDSDISYIRQIFPD
YDGFGLPHALGITSRSQIGVGTTQLPDTLYMYWVSLFDTKFYVTKYEIPDNVKQL
VATKKTYTKWNGFIVDPCYRAEFIFGLLPNGQAKVWLSGCGETIYLTELAPDQIL
DRDSNGVKFDSYRKSSSFADVQQRAKDAGVELEPIPWDKVNKVYSRYEIKTLGE

>HE-25

MATHYCLVENGMEKMLLQWLSEQTSNVYWLADHSTFKQAVANNSGIVFDGTQ
VVFHGETFAPVMALSPWLVPVSDMVSDIDYECLQQGIFLSCSCPSTELLSHLQSL
LIAALEGEEVLFRFYDRQVIVPMLDAMRDLERNDFLGPVEKLAAVKQGVFQEW
GNTRSFEFIYQPAPWWKIQPYHLMPLYRTEVHAQVLERRFWEKLPYVMEQLDEP
QQWIKTILDEAKQANLGHDNAEYLVLNHLWKGSLTTLEQMSDDLHLNQQELQE
IMQIREKLA

>HE-39

MILECRRCGTKKQAAIKELAISASKENKMKKLFFLLLVLLVGSCSTNAAYETEQP
WTWRVGVTMPSFYPVNVTRVYGVNNKEDWTVLIHNFMTFVSDSDISYIRQIFPD
YDGFGLPHALGITSRSQIGVGTTQLPDTLYMYWVSLFDTKFYVTKYEIPDNVKQL
VATKKTYTKWNGFIVDPCYRAEFIFGLLPNGQAKVWLSGCGETIYLTELAPDQIL
DRDSNGVKFDSYRKSSSFADVQQRAKDAGVELEPIPWDKVNKVYSRYEIKTLGE

>HE-45

MILECRRCGTKKQAAIKELAISASKENKMKKLFFLLLVLLVGSCSTNAAYETEQP
WTWRVGVTMPSFYPVNVTRVYGVNNKEDWTVLIHNFMTFVSDSDISYIRQIFPD
YDGFGLPHALGITSRSQIGVGTTQLPDTLYMYWVSLFDTKFYVTKYEIPDNVKQL
VATKKTYTKWNGFIVDPCYRAEFIFGLLPNGQAKVWLSGCGETIYLTELAPDQIL
DRDSNGVKFDSYRKSSSFADVQQRAKDAGVELEPIPWDKVNKVYSRYEIKTLGE

>LMA3984-4

MVIQNFEAFLIAITILTLTPGLDTALVIRNTSRAGFADGCTTSLGICFGLFVHATFS
AIGISAILAQSAELFQIVKMVGAAYLIWLGISSLRSLMKTGQGIEVASLAHAQFRL
TRSLREGFLSNVLNPKTAVFYLAFLPQFINPDYSPLAQSLLMALIHFAIAMVWQC
GLAGALSSAKNLLKNASFMRWMEGTTGVVLVALGIKLLLEKPQA

>M66-2

MKLLNNLAIKKLISVVVLFIALFFLGYDYYQTSQPNIWGDEPDESYITISGKKPIDA
YLEVWTHFWVTGDECEAYSYDLFGQKAHQGGKISQKITHDFAKDGSNFEFRIPY
QTYKNSQNCIVELRDFSIQAHNDFDTVGFAQLRFSPAGREYYNREVDLNSLITAD
NCNSDIFKSIRKEWAGAIGCHFYVDGKKKTKDEEFNAYTIYYDFSKFNNDTVIHY
DILAGENYRSEPLSSAQKTAVVSDDN

>MAK_757

MKLLNNLAIKKLISVVVLFIALFFLGYDYYQTSQPNIWGDEPDESYITISGKKPIDA
YLEVWTHFWVTGDECEAYSYDLFGQKAHQGGKISQKITHDFAKDGSNFEFRIPY
QTYKNSQNCIVELRDFSIQAHNDFDTVGFAQLRFSPAGREYYNREVDLNSLITAD
NCNSDIFKSIRKEWAGAIGCHFYVDGKKKTKDEEFNAYTIYYDFSKFNNDTVIHY
DILAGENYRSEPLSSAQKTAVVSDDN

>MJ-1236

MKLLNNLAIKKLISVVVLFIALFFLGYDYYQTSQPNIWGDEPDESYITISGKKPIDA
YLEVWTHFWVTGDECEAYSYDLFGQKAHQGGKISQKITHDFAKDGSNFEFRIPY
QTYKNSQNCIVELRDFSIQAHNDFDTVGFAQLRFSPAGREYYNREVDLNSLITAD
NCNSDIFKSIRKEWAGAIGCHFYVDGKKKTKDEEFNAYTIYYDFSKFNNDTVIHY
DILAGENYRSEPLSSAQKTAVVSDDN

>MO-10

MKLLNNLAIKKLISVVVLFIALFFLGYDYYQTSQPNIWGDEPDESYITISGKKPIDA
YLEVWTHFWVTGDECEAYSYDLFGQKAHQGGKISQKITHDFAKDGSNFEFRIPY
QTYKNSQNCIVELRDFSIQAHNDFDTVGFAQLRFSPAGREYYNREVDLNSLITAD
NCNSDIFKSIRKEWAGAIGCHFYVDGKKKTKDEEFNAYTIYYDFSKFNNDTVIHY
DILAGENYRSEPLSSAQKTAVVSDDN

>MZO-2

MKLFNNLAIKKLISVVVLFIALFFLGYDYYQTSQPNIWGDEPDESYITISGKKPIDA
YLEVWTHFWVTGDECEAYSYDLFGQKAHQGGKISQKITHDFAKDGSNFEFRIPY
QTYKNSQNCIVELRDFSIQAHNDFDTVGFAQLRFSPAGREYYNREVDLNSLITAD
NCNSDIFKSIRKEWAGAIGCHFYVDGKKKTKDEEFNAYTIYYDFSKFNNDTVIHY
DILAGENYRSEPLSSAQKTAVVSDDN

>MZO-3

MKRLFLILSIVLLSACSSNAASVPGKAWYVGVTMPSFYPVKITQIYGVNEKENWT
VLIHNIIPRMSDSDIEYIRNRFPEYDGFGLPHALGITSRNQIGIGTNHLPDTLYVHW
VSLFDNKFYVTKYDIPDNVKQLSATEVSYTRSDGAVFESCYRTRFVFGLLPNGRA
KVWLSDCGETIYLTELAPDQTPDRDSNGFKADTYKESSYIRNIQQRAKDAGVELE
PIPWDKVNKVYSINEIKTLGE

>N16961

MKLLNNLAIKKLISVVVLFIALFFLGYDYYQTSQPNIWGDEPDESYITISGKKPIDA
YLEVWTHFWVTGDECEAYSYDLFGQKAHQGGKISQKITHDFAKDGSNFEFRIPY
QTYKNSQNCIVELRDFSIQAHNDFDTVGFAQLRFSPAGREYYNREVDLNSLITAD
NCNSDIFKSIRKEWAGAIGCHFYVDGKKKTKDEEFNAYTIYYDFSKFNNDTVIHY
DILAGENYRSEPLSSAQKTAVVSDDN

>NIH41

MKLLNNLAIKKLISVVVLFIALFFLGYDYYQTSQPNIWGDEPDESYITISGKKPIDA
YLEVWTHFWVTGDECEAYSYDLFGQKAHQGGKISQKITHDFAKDGSNFEFRIPY
QTYKNSQNCIVELRDFSIQAHNDFDTVGFAQLRFSPAGREYYNREVDLNSLITAD
NCNSDIFKSIRKEWAGAIGCHFYVDGKKKTKDEEFNAYTIYYDFSKFNNDTVIHY
DILAGENYRSEPLSSAQKTAVVSDDN

>O395

MKLLNNLAIKKLISVVVLFIALFFLGYDYYQTSQPNIWGDEPDESYITISGKKPIDA
YLEVWTHFWVTGDECEAYSYDLFGQKAHQGGKISQKITHDFAKDGSNFEFRIPY
QTYKNSQNCIVELRDFSIQAHNDFDTVGFAQLRFSPAGREYYNREVDLNSLFTAD
NCNSDIFKSIRKEWAGAIGCHFYVDGKKKTKDEEFNAYTIYYDFSKFNNDTVIHY
DILAGENYRSEPLSSAQKTAVVSDDN

>TM11079-80

MKKTLVILSLFLVNACTSNAASVPGEPWYVGVTMPSFYPVRVTQVYGVNNKED
WTVLIHNFMTFVSDSDVEYIRNRFPDYDGFGLPHALGITSRNQIGTGTNHLPDTL
YVYWVSLFDTKFYVTKYDIPDNVKQLSATEVGYTRSDGEFFESCYRTRFVFGLL

PNGQAKVWLSGCGETIYLTELAPDEILDRDSNGVKFDSYRKSSSFADVQQRAKD
AGVELEPIPWDKVNKVYSINEIKTLGE

>TMA21

MKRLFLILSIVLLSACSSNAASFPGKAWYVGVTMPSFYPVKITQIYGVNEKENWT
VLIHNIIPRMSDSDIEYIRNRFPEYDGFGLPHALGITSRNQIGIGTNHLPDTLYVHW
VSLFDNKFYVTKYDIPDNVKQLSATEVSYTRSDGAVFESCYRTRFVFGLLPNGRA
KVWLSDCGETIYLTELAPDQTPDRDSNGFKADTYKESSYISNIQQRAKDAGVELE
PIPWDKVNKVYSINEIKTLGE

>V51

MKKTLVILSLFLVNACTSNAASVPGEPWYVGVTMPSFYPVRVTQVYGVNNKED
WTVLIHNFMTFVSDSDVEYIRNRFPDYDGFGLPHALGITSRNQIGTGTNHLPDTL
YVYWVSLFDTKFYVTKYDIPDNVKQLSATEVGYTRSDGEFFESCYRTRFVFGLL
PNGQAKVWLSGCGETIYLTELAPDEILDRDSNGVKFDSYRKSSSFADVQQRAKD
AGVELEPIPWDKVNKVYSINEIKTLGE

>V52

MKLLNNLAIKKLISVVVLFIALFFLGYDYYQTSQPNIWGDEPDESYITISGKKPIDA
YLEVWTHFWVTGDECEAYSYDLFGQKAHQGGKISQKITHDFAKDGSNFEFRIPY
QTYKNSQNCIVELRDFSIQAHNDFDTVGFAQLRFSPAGREYYNREVDLNSLITAD
NCNSDIFKSIRKEWAGAIGCHFYVDGKKKTKDEEFNAYTIYYDFSKFNNDTVIHY
DILAGENYRSEPLSSAQKTAVVSDDN

>VL426

MILECLRCGTKKKIVIKELAIRASKENKMKKMILLLSVLLLNSCSTNAEYKTEQP
WAWSISVTMPSFYPANVTQVYGVNDKENWTALLHGYLHTMRNSELKRIQERFT
EYDGFGLSLLNTTMGVQIGTGTTHLPDTLYLYWVSLFDTKFYVTKYELPTSVKQ
LIATKTTYTRRDGAIVDSCYRAEFIFGLLPNGQAKVWLDGCGETIYLTELAPDQIL
DRDSNGVKFDSYRKSSSFADVQQRAKDAGVELEPIPWDKVNKVYSINEIKTLGE

Auxiliary cluster 2:

>623-39

MKPYPNRLTRFVLLLSIVVGLMACKPNPQDDIALFQPFITENINKKSADPYISSTV
KPGDKMYEILVNIQHGRHDIASSKLQSLIEGNDSDAMVWYAKLIYRASVNNRPK
ALSLYQKAMDDGNPYAYIMLSPGLISSGCVSYFGEKNCTLEHFNKAIELFKPLAA
QGDLRAQYFLLQQQELDKSKETRAQYIQEVIRFSQAHYYQPLMDYVNTILTRPQ
GKDKYEARTTEQYNLAIDLLTIAANNNYIPAINKLSFLLEDTAQEESERLINLSLK
LGSTSTVKYKYHRSNRGSEEKYFYNALYKGLSGEYSFDVHTLNDEEKAKVNEK
VKLFMKDITPMVYIDGFTSRDDWVD

>1587

MKPYPNRLTHFVLLLSIVVGLMACKPNPQDDIALFQPFITENINKKSADPYISSTV
KPGDKMYEILVNIQHGRHDIASSKLQSLIEGNDSDAMVWYAKLIYRASVNNRPK
ALSLYQKAMDDGNPYAYIMLSPGLISSGCVSYFGEKNCTLEHFNKAIELFKPLAA

QGDLRAQYFLLQQQELDKSKETRAQYIQEVIRFSQAHYYQPLMDYVNTILTRPQ
GKDKYEARTTEQYNLAIDLLTIAANNNYIPAINKLSFLLEDTAQEESERLINLSLK
LGSTSTVKYKYHRSNRGSEEKYFYNALYKGLSGEYSFDVHTLNDEEKAKVNEK
VKLFMKDITPMVYIDGFTSRDDWVD

>2010EL-1786

MLIDKNELEQLKVKLHSSEVIYQWDSVAYGERRSEIFRVFGAISAGIVPLWPFIFF
ADIQFNSKEFWGFICFSLAGMAAARYLFMPDHRYCYSLTQAGIYYTDQEVIPDA
AYTFVRGFAWVGIAVCLLALAVVGPLAFVGAGGFALLAFGLTNFHPTVHKKEV
YFADQLIVFDPIKEKMVDLNTDSTDEPWFDRRLFFSSLDEKTHFIELVKSIHNNVD
YLPLQRVNDQYKHPIFNQELKEE

>2740-80

MLIDKNELEQLKVKLHSSEVIYQWDSVAYGERRSEIFRVFGAISAGIVPLWPFIFF
ADIQFNSKEFWGFICFSLAGMAAARYLFMPDHRYCYSLTQAGIYYTDQEVIPDA
AYTFVRGFAWVGIAVCLLALAVVGPLAFVGAGGFALLAFGLTNFHPTVHKKEV
YFADQLIVFDPIKEKMVDLNTDSTDEPWFDRRLFFSSLDEKTHFIELVKSIHNNVD
YLPLQRVNDQYKHPIFNQELKEE

>12129(1)

MLIDKNELEQLKVKLHSSEVIYQWDSVAYGERRSEIFRVFGAISAGIVPLWPFIFF
ADIQFNSKEFWGFICFSLAGMAAARYLFMPDHRYCYSLTQAGIYYTDQEVIPDA
AYTFVRGFAWVGIAVCLLALAVVGPLAFVGAGGFALLAFGLTNFHPTVHKKEV
YFADQLIVFDPIKEKMVDLNTDSTDEPWFDRRLFFSSLDEKTHFIELVKSIHNNVD
YLPLQRVNDQYKHPIFNQELKEE

>AM-19226

MKPYPNRITRFVMLLSIMIGLVACKPNPEDDIALFQPFITENINKKSDDPYISSTVK
PGDKMYDILVNIQHGKWDVAEGGLLSLIDEGNPDAMLWYALMLLLDNNKRRE
VTNLIFKSLTSGNPYAALAIAKNSHACAYLGSGSLDSQVVQSLGISDPNSARLCT
DNNFKKAIELFKPLAAQGDLRAQYFLLQQQELEKSKETRAQYIQEVIRFSQAHYY
QPLMDYVSTILTRPQDKFKYEAKTPELYQLAIQLLTVAANNNYIPAIEFLVRQDS
DLKSNDDLYLKLKTLGGTHFYKTRFFEFKNEFNNRQLYCYATLYEAITGSDKFFI
GIENEENITADCNIEESMKEMTPMVYIDGFTSRDKWVD

>Amazonia

MIMAYNSEQLTLQPPKEAQGFLRSVLTKGQIFPFTKANETIALSLAGEITPSYIKET
EKEDKLSYWDENSLESETFSNWTQLYLFIVGLAKVALFFFLPLLYLIMISAMLFGP
YGWKDWAGDFGTVTLYTTLPCLLIYGHFKLVSAGHLFLAPFLKSKRVYSLNRQT
GMVTLFKKGNKERFSHPFIEFDCVLMSAPSPQGNLNYNLMLVHRYHDYSVGVPI
GNLIGSNEMVAEYYRLWNMIQRYMDISQPMPDILVLEPARERDPTTAAYDKQTG
RNPRYWRDMSDEEYQQTLKKIADKQKNQPDSGPVLNVFSLS

>BX_330286

MLIDKNELEQLKVKLHSSEVIYQWDSVAYGERRSEIFRVFGAISAGIVPLWPFIFF
ADIQFNSKEFWGFICFSLAGMAAARYLFMPDHRYCYSLTQAGIYYTDQEVIPDA
AYTFVRGFAWVGIAVCLLALAVVGPLAFVGAGGFALLAFGLTNFHPTVHKKEV

YFADQLIVFDPIKEKMVDLNTDSTDEPWFDRRLFFSSLDEKTHFIELVKSIHNNVD
YLPLQRVNDQYKHPIFNQELKEE

>C6706

MLIDKNELEQLKVKLHSSEVIYQWDSVAYGERRSEIFRVFGAISAGIVPLWPFIFF
ADIQFNSKEFWGFICFSLAGMAAARYLFMPDHRYCYSLTQAGIYYTDQEVIPDA
AYTFVRGFAWVGIAVCLLALAVVGPLAFVGAGGFALLAFGLTNFHPTVHKKEV
YFADQLIVFDPIKEKMVDLNTDSTDEPWFDRRLFFSSLDEKTHFIELVKSIHNNVD
YLPLQRVNDQYKHPIFNQELKEE

>CA401

MLIDKNELEQLKVKLHSSEVIYQWDSVAYGERRSEIFRVFGAISAGIVPLWPFIFF
ADIQFNSKEFWGFICFSLAGMAAARYLFMPDHRYCYSLTQAGIYYTDQEVIPDA
AYTFVRGFAWVGIAVCLLALAVVGPLAFVGAGGFALLAFGLTNFHPTVHKKEV
YFADQLIVFDPIKEKMVDLNTDSTDEPWFDRRLFFSSLDEKTHFIELVKSIHNNVD
YLPLQRVNDQYKHPIFNQELKEE

>CIRS101

MLIDKNELEQLKVKLHSSEVIYQWDSVAYGERRSEIFRVFGAISAGIVPLWPFIFF
ADIQFNSKEFWGFICFSLAGMAAARYLFMPDHRYCYSLTQAGIYYTDQEVIPDA
AYTFVRGFAWVGIAVCLLALAVVGPLAFVGAGGFALLAFGLTNFHPTVHKKEV
YFADQLIVFDPIKEKMVDLNTDSTDEPWFDRRLFFSSLDEKTHFIELVKSIHNNVD
YLPLQRVNDQYKHPIFNQELKEE

>CP1041

MLIDKNELEQLKVKLHSSEVIYQWDSVAYGERRSEIFRVFGAISAGIVPLWPFIFF
ADIQFNSKEFWGFICFSLAGMAAARYLFMPDHRYCYSLTQAGIYYTDQEVIPDA
AYTFVRGFAWVGIAVCLLALAVVGPLAFVGAGGFALLAFGLTNFHPTVHKKEV
YFADQLIVFDPIKEKMVDLNTDSTDEPWFDRRLFFSSLDEKTHFIELVKSIHNNVD
YLPLQRVNDQYKHPIFNQELKEE

>DL4211

MVTREINMKKIIVICLMLLHSAAWAMESVEQGIRLFNQKEYQQAQQIFQQQSDA
GSAYATFWLGVAQYKNRQHFEAGQTFLQAAEMGDPWAMGVLSEYEYYVNHP
CGYLGWPCDKKWYDLAIKGWKKEAEKGNGKALLQLQFFTNQWRNYIPFLGRY
LIDKKYIEAFELGSLNAVLYVSRFSNISVEDEIKYLKLAANQGYAPAMETLYYRM
NTIGYDEAMKWINKAIELGYAEAARTLLLSYTLGEKDRDGNIIMPPDPKKAYYY
SRLTEALGGPKQDNSLILYRNVIKDGLPVSDENGEAVLEILVTEQEQADMDKQV
AEFVKDIKPNLFLDETSIDLF

>DL4215

MVTREISMKKFIVICLMLLHSAVWANTPIEQGIRLFNQKEYQQAQQIFQQQSDVG
SAYATFWLGVTQYKNRQHFEAGDTFLKAAEMGDPWAMGVLGDVNLYANNPC
KFLGWPCDEKWLTKAKQGWKVLAENGNGKAAFALKINQREWWEYIPFYRQSR
YQEIVSKAIPNGGYKFLDYNTYWDSSEAKLPYLKLAANQGYAPAMETLYYRMD
TIGYDEAMKWINKAIELGYAEAARTLYLAYRVGEKDRDGNVILQPDPKKAYFY

NRLTGALGGEEKLAHLITQEPVHDDDGIPLADENGEPVFEILVTEQEQAEMDKQV
AEFVKDIKPNLFLDETSIDLF

>HC-07A1

MLIDKNELEQLKVKLHSSEVIYQWDSVAYGERRSEIFRVFGAISAGIVPLWPFIFF
ADIQFNSKEFWGFICFSLAGMAAARYLFMPDHRYCYSLTQAGIYYTDQEVIPDA
AYTFVRGFAWVGIAVCLLALAVVGPLAFVGAGGFALLAFGLTNFHPTVHKKEV
YFADQLIVFDPIKEKMVDLNTDSTDEPWFDRRLFFSSLDEKTHFIELVKSIHNNVD
YLPLQRVNDQYKHPIFNQELKEE

>HC-32A1

MLIDKNELEQLKVKLHSSEVIYQWDSVAYGERRSEIFRVFGAISAGIVPLWPFIFF
ADIQFNSKEFWGFICFSLAGMAAARYLFMPDHRYCYSLTQAGIYYTDQEVIPDA
AYTFVRGFAWVGIAVCLLALAVVGPLAFVGAGGFALLAFGLTNFHPTVHKKEV
YFADQLIVFDPIKEKMVDLNTDSTDEPWFDRRLFFSSLDEKTHFIELVKSIHNNVD
YLPLQRVNDQYKHPIFNQELKEE

>HC-33A2

MLIDKNELEQLKVKLHSSEVIYQWDSVAYGERRSEIFRVFGAISAGIVPLWPFIFF
ADIQFNSKEFWGFICFSLAGMAAARYLFMPDHRYCYSLTQAGIYYTDQEVIPDA
AYTFVRGFAWVGIAVCLLALAVVGPLAFVGAGGFALLAFGLTNFHPTVHKKEV
YFADQLIVFDPIKEKMVDLNTDSTDEPWFDRRLFFSSLDEKTHFIELVKSIHNNVD
YLPLQRVNDQYKHPIFNQELKEE

>HC-38A1

MLIDKNELEQLKVKLHSSEVIYQWDSVAYGERRSEIFRVFGAISAGIVPLWPFIFF
ADIQFNSKEFWGFICFSLAGMAAARYLFMPDHRYCYSLTQAGIYYTDQEVIPDA
AYTFVRGFAWVGIAVCLLALAVVGPLAFVGAGGFALLAFGLTNFHPTVHKKEV
YFADQLIVFDPIKEKMVDLNTDSTDEPWFDRRLFFSSLDEKTHFIELVKSIHNNVD
YLPLQRVNDQYKHPIFNQELKEE

>HC-43B1

MLIDKDELEQLKANLHSSEVLYQWECVAFSERRGDIARWIIAVIAGFLLPSLFLVF
GDIAFNSTEFWGFISFGLAGMSAGRYLLMPDHRYCYSLTQAGIYYTDQEVIPDAA
YTFVRGFAWVGIAVCLLALAVVGPLAFVGAGGFALLAFGLTNFHPTVHQKDVY
FADQLILFDPIKEKMVKLNSESRRHPRFSRTLFFSSFDEKTRFIELVKSIHNNVDYL
PLQRVNDQYKHPIFNQELKEE

>HE-25

MKPYPNRIARFVLLLSIVVGLMACKPNPQDDISLFQPFITENINKKSADPYISSTVK
PGDKMYDILVNIQHGDTKPAYEKLMVLVDKNDYEAQFWVAKMIYQASSQSIAR
AINLYKQSSNQGNPYAYYELSPLGKGCLAYFGKKTCTDENLNKAIELFKPLAAQ
GDLRAQYFLLQQQELDKSKETRAQYIHELIRFSKNHYYQPLMDYVNTILTRPQG
KFKYEAKTPELYQLAIQLLTVAANNNYIPAIEFLVRQDSDLKSNDDLYLKLKTLG
GTHFYKTRFFEFKNEFNNRQLYCYATLYEAITGSDKFFIGIENEENIAADCNIEDS
MKEMTPMVYIDGFTSRDDWVD

>HE-39

MKPYPNRLTRFVLLLSIVVGLMACKPNPQDDIALFQPFITENINKKSADPYISSTV
KPGDKMYEILVNIQHGRHDIASSKLQSLIEGNDSDAMVWYAKLIYRASVNNRPK
ALSLYQKAMDDGNPYAYIMLSPGLISSGCVSYFGEKNCTLEHFNKAIELFKPLAA
QGDLRAQYFLLQQQELDKSKETRAQYIQEVIRFSQAHYYQPLMDYVNTILTRPQ
GKDKYEARTTEQYNLAIDLLTIAANNNYIPAINKLSFLLEDTAQEESERLINLSLK
LGSTSTVKYKYHRSNRGSEEKYFYNALYKGLSGEYSFDVHTLNDEEKAKVNEK
VKLFMKDITPMVYIDGFTSRDDWVD

>HE-45

MIMAYNSEQLTLQPPKEAQGFLRSVLTKGQIFPFTKANETIALSLAGEITPSYIKET
EKEDKLSYWDENSLESETFSNWTQLYLFIVGLAKVALFFFLPLLYLIMISAMLFGP
YGWKDWAGDFGTVTLYTTLPCLLIYGHFKLVSAGHLFLAPFLKSKRVYSLNRQT
GMVTLFKKGNKERFSHPFIEFDCVLMSAPSPQGNLNYNLMLVHRYHDYSVGVPI
GNLIGSNEMVAEYYRLWNMIQRYMDISQPMPDILVLEPARERDPTTAAYDKQTG
RNPRYWRDMSDEEYQQTLKKIADKQKNQPDSGPVLNVFSLS

>LMA3984-4

MLIDKDELEQLKANLHSSEVLYQWECVAFSERRGDIARWIIAVIAGFLLPSLFLVF
GDIAFNSTEFWGFISFGLAGMSAGRYLLMPDHRYCYSLTQAGVHYTDQEVIPDA
AYTFVRGFAWVGIAVCLLALAVVGPLAFVGAGGFALLAFGLTNFRPTVKHEDV
YFAEQLILFDPIKEKMVDFEH

>M66-2

MLIDKNELEQLKVKLHSSEVIYQWDSVAYGERRSEIFRVFGAISAGIVPLWPFIFF
ADIQFNSKEFWGFICFSLAGMAAARYLFMPDHRYCYSLTQAGIYYTDQEVIPDA
AYTFVRGFAWVGIAVCLLALAVVGPLAFVGAGGFALLAFGLTNFHPTVHKKEV
YFADQLIVFDPIKEKMVDLNTDSTDEPWFDRRLFFSSLDEKTHFIELVKSIHNNVD
YLPLQRVNDQYKHPIFNQELKEE

>MAK_757

MLIDKNELEQLKVKLHSSEVIYQWDSVAYGERRSEIFRVFGAISAGIVPLWPFIFF
ADIQFNSKEFWGFICFSLAGMAAARYLFMPDHRYCYSLTQAGIYYTDQEVIPDA
AYTFVRGFAWVGIAVCLLALAVVGPLAFVGAGGFALLAFGLTNFHPTVHKKEV
YFADQLIVFDPIKEKMVDLNTDSTDEPWFDRRLFFSSLDEKTHFIELVKSIHNNVD
YLPLQRVNDQYKHPIFNQELKEE

>MJ-1236

MLIDKNELEQLKVKLHSSEVIYQWDSVAYGERRSEIFRVFGAISAGIVPLWPFIFF
ADIQFNSKEFWGFICFSLAGMAAARYLFMPDHRYCYSLTQAGIYYTDQEVIPDA
AYTFVRGFAWVGIAVCLLALAVVGPLAFVGAGGFALLAFGLTNFHPTVHKKEV
YFADQLIVFDPIKEKMVDLNTDSTDEPWFDRRLFFSSLDEKTHFIELVKSIHNNVD
YLPLQRVNDQYKHPIFNQELKEE

>MO-10

MLIDKNELEQLKVKLHSSEVIYQWDSVAYGERRSEIFRVFGAISAGIVPLWPFIFF
ADIQFNSKEFWGFICFSLAGMAAARYLFMPDHRYCYSLTQAGIYYTDQEVIPDA
AYTFVRGFAWVGIAVCLLALAVVGPLAFVGAGGFALLAFGLTNFHPTVHKKEV

YFADQLIVFDPIKEKMVDLNTDSTDEPWFDRRLFFSSLDEKTHFIELVKSIHNNVD
YLPLQRVNDQYKHPIFNQELKEE

>MZO-2

MLIDKNELEQLKVKLHSSEVIYQWDSVAYGERRSEIFRVFGAISAGIVPLWPFIFF
ADIQFNSKEFWGFICFSLAGMAAARYLFMPDHRYCYSLTQAGIYYTDQEVIPDA
AYTFVRGFAWVGIAVCLLALAVVGPLAFVGAGGFALLAFGLTNFHPTVHKKEV
YFADQLIVFDPIKEKMVDLNTDSTDEPWFDRRLFFSSLDEKTHFIELVKSIHNNVD
YLPLQRVNDQYKHPIFNQELKEE

>MZO-3

MVTREISMKKFIVICLMLLHSAVWANTPIEQGIRLFNQKEYQQAQQIFQQQSDVG
SAYATFWLGVTQYKNRQHFEAGDTFLKAAEMGDPWAMGVLGDVNLYANNPC
KFLGWPCDEKWLTKAKQGWKVLAENGNGKAAFALKINQREWWEYIPFYRQSR
YQEIVSKAIPNGGYKFLDYNTYWDSSEAKLPYLKLAANQGYAPAMETLYYRMD
TIGYDEAMKWINKAIELGYAEAARTLYLAYRVGEKDRDGNVILQPDPKKAYFY
NRLTGALGGEEKLAHLITQEPVHDDDGIPLADENGEPVFEILVTEQEQAEMDKQV
AEFVKDIKPNLFLDETSIDLF

>N16961

MLIDKNELEQLKVKLHSSEVIYQWDSVAYGERRSEIFRVFGAISAGIVPLWPFIFF
ADIQFNSKEFWGFICFSLAGMAAARYLFMPDHRYCYSLTQAGIYYTDQEVIPDA
AYTFVRGFAWVGIAVCLLALAVVGPLAFVGAGGFALLAFGLTNFHPTVHKKEV
YFADQLIVFDPIKEKMVDLNTDSTDEPWFDRRLFFSSLDEKTHFIELVKSIHNNVD
YLPLQRVNDQYKHPIFNQELKEE

>NIH41

MLIDKNELEQLKVKLHSSEVIYQWDSVAYGERRSEIFRVFGAISAGIVPLWPFIFF
ADIQFNSKEFWGFICFSLAGMAAARYLFMPDHRYCYSLTQAGIYYTDQEVIPDA
AYTFVRGFAWVGIAVCLLALAVVGPLAFVGAGGFALLAFGLTNFHPTVHKKEV
YFADQLIVFDPIKEKMVDLNTDSTDEPWFDRRLFFSSLDEKTHFIELVKSIHNNVD
YLPLQRVNDQYKHPIFNQELKEE

>O395

MLIDKNELEQLKVKLHSSEVIYQWDSVAYGERRSEIFRVFGAISAGIVPLWPFIFF
ADIQFNSKEFWGFICFSLAGMAAARYLFMPDHRYCYSLTQAGIYYTDQEVIPDA
AYTFVRGFAWVGIAVCLLALAVVGPLAFVGAGGFALLAFGLTNFHPTVHKKEV
YFADQLIVFDPIKEKMVDLNTDSTDEPWFDRRLFFSSLDEKTHFIELVKSIHNNVD
YLPLQRVNDQYKHPIFNQELKEE

>TM11079-80

MIMAYNSEQLTLQPPKEAQGFLRSVLTKGQIFPFTKANETIALSLAGEITPSYIKET
EKEDKLSYWDENSLESETFSNWTQLYLFIVGLAKVALFFFLPLLYLIMISAMLFGP
YGWKDWAGDFGTVTLYTTLPCLLIYGHFKLVSAGHLFLAPFLKSKRVYSLNRQT
GMVTLFKKGNKERFSHPFIEFDCVLMSAPSPQGNLNYNLMLVHRYHDYSVGVPI
GNLIGSNEMVAEYYRLWNMIQRYMDISQPMPDILVLEPARERDPTTAAYDKQTG
RNPRYWRDMSDEEYQQTLKKIADKQKNQPDSGPVLNVFSLS

>TMA21

MESDIIMFWKKKSQDIDLKELVPKAGNTRKFLTGEAVFFSPKPLPSAKIASETIVQ
SPEILEINDTYLDIGQRNQGKAWQVQIMSSGISVSAIVILLILFYVGHRDMQLFGSF
FIAFLESFNDVAVSLTIAFFLLLCVSCQVIIGTSIDFARQRPIRLNRQRREICYYSDR
EKKSIIAPWEEVVCWVALNRGTTGNSMVTHFTFGLAIPTPDGKDYWTLRRPIASV
TDGQRMWETMRMYMDEPLTRRPVPSPFIKEDRAYFDKSRREMCDRFQKGPKC
WFILNSNAPWVSYASMFFYYLFHILSGWKLPYWVSEWTDNLAKVDLPKEVEEW
SKPIPESEWAKPSEELLRQRAAIEKHYEAGGGIMDFNPATQ

>V51

MKPYPNLITRFVLLLSIMIGLVACKPNPQDDIALFQPFITENINKKSADPYISSTVKP
GDKMYEILVNIQHGRHDIASSKLQSLIEVNDSDAMVWYAKLIYRASVNNRTKAL
SLYQKAMDNDNPYAYIMLSPGLISSGCVSYFGEKNCTLEHFNKAIELFKPLAAQG
DVRAQYFLLQQQELDKSKETRAQYIQEVIRFSQAHYYQPLMDYVNTILTRPQGK
DKYEARTTEQYNLAIDLLTIASNYNYVPAIKRLLEVDLDEEVFTKRMEQLTRLGS
TFEIYENYFNSQDEHERHYYSMLFKQLTGDNKLSSSYSATDDEVKMINEKIEKVI
SSMTPMVYIDGFTSRDNWVD

>V52

MLIDKNELEQLKVKLHSSEVIYQWDSVAYGERRSEIFRVFGAISAGIVPLWPFIFF
ADIQFNSKEFWGFICFSLAGMAAARYLFMPDHRYCYSLTQAGIYYTDQEVIPDA
AYTFVRGFAWVGIAVCLLALAVVGPLAFVGAGGFALLAFGLTNFHPTVHKKEV
YFADQLIVFDPIKEKMVDLNTDSTDEPWFDRRLFFSSLDEKTHFIELVKSIHNNVD
YLPLQRVNDQYKHPIFNQELKEE

>VL426

MLIDKDELEQLKANLHSSEVLYQWECVAFSERRGDIARWVIGIFAGFLPALLFIPL
GDIPFNSKGFWGFMIFGIVFALVARYLLMPDHRYCYSLTQAGVHYTDQEVIPDA
AYTFVRGFAWVGIAVCLLALAVVGPLAFVGAGGFALLAFGLTKFHPTVHQKEF
YFADQLIVFDPIKEKMLIFNTELVRDPIFIFSFMGTLYFNSFDEKNRVMALIKNVH
NNIDYCPLQRVNDEFKHRKYQQALKDE

Large cluster:

>623-39_

MNIFKKIIYCIILIMFSLQAVGGSIPSSSQLRSLLRSYDIQTANIGETIAYIPIEKPCLN
LLARNEHYEYCVINEGENLSNGDKEGFFITLDSVSESSVSFSYNTIWYSVNCRLSI
SDKKIKCDKAE

>1587_

MNIFKKIIYCIILIMFSLQAVGGSIPSSSQLRSLLRSYDIQTANIGETIAYIPIEKPCLN
LLARNEHYEYCVINEGENLSNGDKEGFFITLDSVSESSVSFSYNTIWYSVNCRLSI
SDKKIKCDKAE

>2010EL-1786

MNNLLSAYVTMLLILLSISGGAIASENCNDTSGVHQKILVCIQNEIAKSETQIRNNI
SSKSIDYGFPDDFYSKQRLAIHEKCMLYINVGGQRGELLMNQCELSMLQGLDIYI
QQYIEDVDNS

>2740-80

MNNLLSAYVTMLLILLSISGGAIASENCNDTSGVHQKILVCIQNEIAKSETQIRNNI
SSKSIDYGFPDDFYSKQRLAIHEKCMLYINVGGQRGELLMNQCELSMLQGLDIYI
QQYIEDVDNS

>12129(1)

MNIFKKIIYCIILIMFSLQAVGGSIPSSSQLRSLLRSYDIQTANIGETIAYIPIEKPCLN
LLARNEHYEYCVINEGENLSNGDKEGFFITLDSVSESSVSFSYNTIWYSVNCRLSI
SDKKIKCDKAE

>AM19226_

MLKYLFFIGFIFSSSASASASDECDYTLPLGSIVACYENKNEESDAVLNKSYNDLK
KLVMGSPYDDETKGVYWSNIVKSQKNWIQMRDSQCLAKGVFFENGTDLQRIEIK
KCLFLSTEIRVLYLNEEILFIKNLP

>Amazonia

MKNCLYLIVLTLCFSIARANDENVDDLIVSSYIETQDEKIKDFQGKEFQISFQKESI
LKKEGLNDGSLKAFIRISGSQSLLFEQKDFKPYGQDVFVFSCKASDANMRVFIQT
RYASSDNITPWFVQFTPYIFTIESGSVRREYKFENEYFYGGDFFYRRIIDDEKLIIKY
PFYTNDKFIEKLFESKLCVNYSE

>BX330286_

MNNLLSAYVTMLLILLSISGGAIASENCNDTSGVHQKILVCIQNEIAKSETQIRNNI
SSKSIDYGFPDDFYSKQRLAIHEKCMLYINVGGQRGELLMNQCELSMLQGLDIYI
QQYIEDVDNS

>C6706_

MNNLLSAYVTMLLILLSISGGAIASENCNDTSGVHQKILVCIQNEIAKSETQIRNNI
SSKSIDYGFPDDFYSKQRLAIHEKCMLYINVGGQRGELLMNQCELSMLQGLDIYI
QQYIEDVDNS

>CA401

MNNLLSAYVTMLLILLSISGGAIASENCNDTSGVHQKILVCIQNEIAKSETQIRNNI
SSKSIDYGFPDDFYSKQRLAIHEKCMLYINVGGQRGELLMNQCELSMLQGLDIYI
QQYIEDVDNS

>CIRS101

MNNLLSAYVTMLLILLSISGGAIASENCNDTSGVHQKILVCIQNEIAKSETQIRNNI
SSKSIDYGFPDDFYSKQRLAIHEKCMLYINVGGQRGELLMNQCELSMLQGLDIYI
QQYIEDVDNS

>CP1041

MNNLLSAYVTMLLILLSISGGAIASENCNDTSGVHQKILVCIQNEIAKSETQIRNNI
SSKSIDYGFPDDFYSKQRLAIHEKCMLYINVGGQRGELLMNQCELSMLQGLDIYI
QQYIEDVDNS

>DL4211_

MVRLFGFLFFLSFSILAKDLDNSQVYSVPSDERWLLIDILKNPCDVCTSDIYIQSGS
ARINGVWVSGVFEFSVSEKIDILFDSDTKFALGDIVQKISVETESNENM

>DL4215_

MNIFKKIIYCIILIMFSLQAVGGSIPSSSQLRSLLRSYDIQTANIGETIAYIPIEKPCLN
LLARNEHYEYCVINEGENLSNGDKEGFFITLDSVSESSVSFSYNTIWYSVNCRLSI
SDKKIKCDKSE

>HC-07A1

MNNLLSAYVTMLLILLSISGGAIASENCNDTSGVHQKILVCIQNEIAKSETQIRNNI
SSKSIDYGFPDDFYSKQRLAIHEKCMLYINVGGQRGELLMNQCELSMLQGLDIYI
QQYIEDVDNS

>HC-32A1_

MNNLLSAYVTMLLILLSISGGAIASENCNDTSGVHQKILVCIQNEIAKSETQIRNNI
SSKSIDYGFPDDFYSKQRLAIHEKCMLYINVGGQRGELLMNQCELSMLQGLDIYI
QQYIEDVDNS

>HC-33A2

MNNLLSAYVTMLLILLSISGGAIASENCNDTSGVHQKILVCIQNEIAKSETQIRNNI
SSKSIDYGFPDDFYSKQRLAIHEKCMLYINVGGQRGELLMNQCELSMLQGLDIYI
QQYIEDVDNS

>HC-38A1

MNNLLSAYVTMLLILLSISGGAIASENCNDTSGVHQKILVCIQNEIAKSETQIRNNI
SSKSIDYGFPDDFYSKQRLAIHEKCMLYINVGGQRGELLMNQCELSMLQGLDIYI
QQYIEDVDNS

>HC-43B1_

MKRLIILLCVFFSFGCNAKAIPDKEELIFFARSYSDFESGSVVVYVPVNKPCINVIS
SGNNYEFCDLQKNINGVNATLDKGAESGIWVVGLKLDISMVEFTYRTPYFDEVC
TIDLLNDNKLQCGNGK

>HE-25

MLKYLFFIGFIFSSSASASDECDYTLPLGSIVACYENKNEESDAVLNKSYNDLKKL
VMGSPYDDETKGVYWSNIVKSQKNWIQMRDSQCLAKGVFFENGTDLQRIEIKK
CLFLSTEIRVLYLNEEILFIKNLP

>HE-39_

MNNLLSAYVTILLILLSISGGAIASENCNDTSGVHQKILVCIQNEIAKSETQIRNNIS
SKSIDYGFPDDFYSKQRLAIHEKCMLYINVGGQRGELLMNQCELSMLQGLDIYIQ
QYIEDVDNS

>HE-45_

MKRLIILLCVFFSFGCNAKAIPDKEELIFFARSYSDFESGSVVVYVPVNKPCINVIS
SGNNYEFCDLQKNINGVNATLDKGAESGIWVVGLKLDISMVEFTYRTPYFDEVC
TIDLLNDNKLQCGNGK

>LMA3984-4

MNIFKKIIYCIILIMFSLQAVGGSIPSSSQLRSLLRSYDIQTANIGETIAYIPIEKPCLI
NEGFFITLDSVSESSVGFFITLDSVSESSVSFSYNTIWYSVNCRLSISDKKIKCDKAE

>M66-2

MNNLLSAYVTMLLILLSISGGAIASENCNDTSGVHQKILVCIQNEIAKSETQIRNNI
SSKSIDYGFPDDFYSKQRLAIHEKCMLYINVGGQRGELLMNQCELSMLQGLDIYI
QQYIEDVDNS

>MAK_757

MNNLLSAYVTMLLILLSISGGAIASENCNDTSGVHQKILVCIQNEIAKSETQIRNNI
SSKSIDYGFPDDFYSKQRLAIHEKCMLYINVGGQRGELLMNQCELSMLQGLDIYI
QQYIEDVDNS

>MJ1236

MNNLLSAYVTMLLILLSISGGAIASENCNDTSGVHQKILVCIQNEIAKSETQIRNNI
SSKSIDYGFPDDFYSKQRLAIHEKCMLYINVGGQRGELLMNQCELSMLQGLDIYI
QQYIEDVDNS

>MO10_

MNNLLSAYVTMLLILLSISGGAIASENCNDTSGVHQKILVCIQNEIAKSETQIRNNI
SSKSIDYGFPDDFYSKQRLAIHEKCMLYINVGGQRGELLMNQCELSMLQGLDIYI
QQYIEDVDNS

>MZO-2_

MNKNFEQSRRDPLLTEDTFHTKTIDYFEATLNDEELISIISDQTKEKDHRKRVSWS
VFGELFEKTILDYSAGKDFVEIQSILHKTLSFYEEHKNNFSEHSLKYWEPDSYQYI
LWLFSLNTLLGDSSFSLSLIRCVSVSGNDDKLLSILFSRLGIIGFPRGEELIFPKTYQ
YLFDAIKGDGAYPSKEERQKSIKKYLSAWYKSMKNCYWHDRHKGRFPTFFGYW
AFEAALVTVLYDLDDSSYRDMLHYPKDLVDHARNNGVDKLFHTQNIKQRWIVF
PNDESPISGRWFDNLSNEVISTVKGERMPGVFENAQGYRHFWVSE

>MZO-3

MNNLLSAYVTMLLILLSISGGAIASENCNDTSGVHQKILVCIQNEIAKSETQIRNNI
SSKSIDYGFPDDFYSKQRLAIHEKCMLYINVGGQRGELLMNQCELSMLQGLDIYI
QQYIEDVDNS

>N16961

MNNLLSAYVTMLLILLSISGGAIASENCNDTSGVHQKILVCIQNEIAKSETQIRNNI
SSKSIDYGFPDDFYSKQRLAIHEKCMLYINVGGQRGELLMNQCELSMLQGLDIYI
QQYIEDVDNS

>NIH41

MNNLLSAYVTMLLILLSISGGAIASENCNDTSGVHQKILVCIQNEIAKSETQIRNNI
SSKSIDYGFPDDFYSKQRLAIHEKCMLYINVGGQRGELLMNQCELSMLQGLDIYI
QQYIEDVDNS

>O395_

MNNLLSAYVTMLLILLSISGGAIASENCNDTSGVHQKILVCIQNEIAKSETQIRNNI
SSKSIDYGFPDDFYSKQRLAIHEKCMLYINVGGQRGELLMNQCELSMLQGLDIYI
QQYIEDVDNS

>TM11079-80_

MKNCLYLIVLTLCFSIARANDENVDDLIVSSYIETQDEKIKDFQGKEFQISFQKESI
LKKEGLNDGSLKAFIRISGSQSLLFEQKDFKPYGQDVFVFSCKASDANMRVFIQT
RYASSDNITPWFVQFTPYIFTIESGSVRREYKFENEYFYGGDFFYRRIIDDEKLIIKY
PFYTNDKFIEKLFESKLCVNYSE

>TMA21_

MLKYLFFIGFIFSSSANANASDECDYTLPLGSIVACYENKNEESDAVLNKSYNDL
KKLVMGSPYDDETKGVYWSNIVKSQKNWIQMRDSQCLAKGVFFENGTDLQRIEI
KKCLFLSTEIRVLYLNEEILFIKNLP

>V51_

MNNLLSAYVTMLLILLSISGGAIASENCNDTSGVHQKILVCIQNEIAKSETQIRNNI
SSKSIDYGFPDDFYSKQRLAIHEKCMLYINVGGQRGELLMNQCELSMLQGLDIYI
QQYIEDVDNS

>V52_

MNNLLSAYVTMLLILLSISGGAIASENCNDTSGVHQKILVCIQNEIAKSETQIRNNI
SSKSIDYGFPDDFYSKQRLAIHEKCMLYINVGGQRGELLMNQCELSMLQGLDIYI
QQYIEDVDNS

>VL426_

MNNLLSAYVTMLLILLSISGGAIASENCNDTSGVHQKILVCIQNEIAKSETQIRNNI
SSKSIDYGFPDDFYSKQRLAIHEKCMLYINVGGQRGELLMNQCELSMLQGLDIYI
QQYIEDVDNS

Sequences of genes from genomes that were not annotated at all:

Amazonia, large cluster, effector-encoding gene

ATGGCAAGGTTACAGTTTCAATTAAAGGTGGATGGCCTTGAGGATGAATCCC
TCGTCGTACGCGGATTTGAAGGGCAAGAGTCCTTGTCTGATTCTGTATGGCGC
AGTGAACCCTGCTATGGTTTTCGCTATCAGGTTGATTTAGCCAGTGCGCTAAG
TAATCTCACGGCTGAGCAATTTGTAGACCAAACCGCACATTTAACCATTCTGC
GTGACGGGCAAGTAGTGCAGCAAATAAACGGTATTGTGCGTCAATTCAGCAA
AGGGGATACCGGCCATCGACATACTTTCTATTCATTAACGCTAGTTCCTGCGC
TAGAACGGCTTTCTTTACGCAGTAATAGCCGTATTTTCCAGCAGCAAAGCGTG
CCAGAGATTATCTCTATTTTACTGCAAGAGATGGGGATTGAAGATTACGCCTT
TGCCCTTAAGCGTGAGTGTGCCCAGCGAGAGTTTTGTGTACAGTATCGTGAA
ACGGATTTGCAGTTTTTACACCGTATCGCCGCCGAAGAAGGCTTGGTATACA
GTCACCTACATGAAGCGCAGAAACACACCTTGCTGTTTACCGATAGCTCGGA
CAGCCAACCGAAGTTAGCCAAGCCAGTACCTTATAATGCGCTGGCGGGTGGC
GAGATCAATCTCCCTTATGTGGTCGATCTGCAATTCAAGACCACCGCACAAGT
CAGCCATACCGAACTTAAAGATTACAGCTTTAAAAAGCCAGCTTACGGATTT
ACTCAGCGCACGCAGGGTAAAGATATCGCTTATCAGCAGCCAAATTATGAGC
ATTTTGATGCTCCGGGACGCTACAAAGACGATGCAAATGGCAAGGCCTTTAG
CCAAATTCGCTTAGAGTATTTACGTCGCGATGCTTTGCTTGCTGATGCGAAGA
GTGACGAACCTTTGCTGTTGGCGGGAGTGCGTTTTGACTTGCAAGATCACCTT
GATCCCGCGATGAATCGAGATTGGCTCGTGGTACAAGCCAATCACCAAGGGA
CTCAGCCACAAGCGTTACAAGAAGAAGGTGGTTCAGGCGCTACCACTTACAG
CAATCAGCTAAAACTTATCCCCGCTCACATCACTTGGCGAGCAAGGCCTTGT
GCTAAGCCGCAAGTGGATGGTCCTATGATTGCCACCGTAGTTGGGCCACAAG
GTGAAGAGATTTATTGCGATAACTTTGGTCGCGTGAAAGTGCATTTCCCGTGG
GATCGCTACTCAAGTAGCAATGAGAAAAGCTCTTGCTGGGTGCGAGTGGCAC
AAGAATGGGCGGGTAGCCAATACGGTAGTATGGCGATTCCGCGAGTTGGCCA
TGAAGTGATTGTTTCGTTCCTGAATGGCGATCCAGATCAACCCATCATCACGG
GGCGTACGTATCATGCGACCAATACCGCACCTTACGCCTTGCCTGACCACAA
AACCAAAACCGTACTGCGCACTGAAACCCACCAAGGACAAGGCTACAACGA
ACTGAGTTTTGAGGATCAAGCGGGCAGCGAACAGATTTTGCTGCATGCACAA
AAAGATTGGGATGCGTTGATTGAGCATGATCACACCGAAGTGATTCGCCACG
ATCAGCATCTCACTGTGGACAATGATCGCTTTACCCGCATTCAGCGCAACCA
ACATTTGACGGTTGAAGGTGAAGTCCGTAGCAAAATCGCACTCGATAGCAGC
CATGAAGTCGGCGCGTCACTACAACACAAAGTCGGGCAACGCATTGCTGTCG
AGGCGGGTAAAGAGATTTCACTCAAAAGCGGCGCAAAAATCGTGGTTGAAG
CCGGAGCGGAGTTAACCCTAAAAGCAGGAGGCAGCTTTGTAAAAGTCGATGC
AGGTGGCGTACATCTAGTCGGTCCTGCCATTAACCTAAACGCGGGCGGCAGT
GCAGGCAGCGGTAGCGCTTATGGTGGGCAATTAGCGGCTGCACCAAGAATGT
TAGCGCAAGCTAAACCAGTAGCAGAATTGGTTCAGCCGGATATTGCTGCCTC
AATGCAATCAGGCGCGGCGCGTGTTATTGATGTGGCATCACTCCCCACAATG
ATACCGAGTTCAGCCAATAACACGGCAAATGATGAGCCTGTGGCTGAAGAGA
AAACACCGGAGCGAATTTTAAAATCCGATCTTCTCAAACCATCCGATGAGTT
AGAGAAACTCGCCAAGCGACAAGCGAGTGCTTATCGTCAAGGTAATCACAGC

GATGAGGTTAAGTTACTACAAGAAGCGCTCATAAAACTGGGGTTTGATCTCG
GGAAAGCAGGAGCTGATGGTGACTTTGGGAGCAAAACTAAAACCGCGATTG
AGCAGTTCCAGAAAAGCTATCAACCAAGCCACCAAACCCACCCGTCTTACAG
CATTGGCGCTGTTGACGGTATTGTCGGCAAAGGTACACTACTGGCACTAGAC
GAAGCCTTGATGGATGGGTGGGTTTATCTTCTTATTACTAAGGATATGCTAAA
TGTAGTGTGCCCTAATTCAGGTAATAATAAGGAACTGTTAGATTCATTAAATA
AATACTGTCCTATGTACGAAATAGATACACCCATTAGAATTGCACATTTTTTA
TCTCAAATTGCTCATGAAAGTGGATGTTTTAGTAGTTTGACAGAAGGAAGTA
ATTATAGTCATATGGCTGCGAAACTAAAGTTTTCAAAATATAGAACTTATGTT
GACTCATTAAAGTCTGGTTCGACAAACAAGTTCATAAGAGAAGATAATCCAG
CAAGATGCAAACAACCTGATTTGTTCAATTTTGTTTATTCAAGTAGTAATGGT
AATGGAGATGAGAAATCTGGAGATGGATATAAATATAGAGGTAGGGGGCTC
ATTCAGATTACAGGAAAAGATAAGTATTCAAGGTTTACAAGTGTGCATAATA
GTAAAAATCCGGATGACACTCAAAATTTTGTCATAAATCCAGATCTTGTTTCC
GAAAATTTGAATTATGCAGTTGAGTCAGCATGTATTTACTGGAGGCATTGGG
GAGCACTATCTAAAAAGTTTAACGCTAATGGTGATATAAATATATTAATTGAT
AATGCACCAAATGATGTTGAGTTAATTAGTCAGGCGGTAAATGGTGGTAGTT
ATGGACACTCTAATGGCTTAGATGACCGCATTGATAAATTTAATAAAATAAA
AAATAACTTGGGATTGTGA

Amazonia, large cluster, immunity protein-encoding gene

ATGAAAAATTGTTTGTATCTTATCGTCTTAACTCTATGTTTCAGTATAGCTCGA
GCTAATGATGAAAATGTTGATGATTTAATTGTATCATCATATATAGAAACTCA
AGATGAGAAGATAAAAGATTTCCAAGGAAAAGAGTTTCAAATTTCTTTCCAG
AAAGAGAGTATCCTGAAAAAAGAAGGTCTCAATGATGGTAGCCTTAAAGCAT
TTATCAGAATCTCTGGTAGTCAAAGTTTATTATTTGAACAAAAAGATTTCAAG
CCATATGGACAAGATGTATTTGTTTTTTCTTGTAAGGCATCAGATGCCAACAT
GAGAGTTTTTATCCAGACGAGATATGCAAGTTCTGATAATATAACACCTTGGT
TTGTTCAGTTTACGCCATATATTTTTACGATTGAATCTGGTTCAGTAAGGAGA
GAATATAAATTTGAAAATGAATATTTCTATGGTGGCGATTTTTTTTATAGAAG
AATAATTGATGATGAGAAATTAATCATTAAATATCCATTCTATACAAATGAC
AAATTCATAGAGAAATTATTTGAATCTAAATTATGTGTTAATTATAGTGAATA
A

C6706, large cluster, effector-encoding gene

ATGGCAAGGTTACAGTTTCAATTAAAGGTGGATGGCCTTGAGGATGAATCCC
TCGTCGTACGCGGATTTGAAGGGCAAGAGTCCTTGTCTGATTCTGTATGGCGC
TGTGAACCCTGCTATGGTTTTCGCTATCAGGTTGATTTAGCCAGTGCGCTAAG
CAACCTTACTGCTGAGCAATTTGTCGACCAAACCGCGCATTTAACCATTCTGC
GTGACGGGCAAGTGGTGCAGCAAATTAACGGTATCGTACGCCAATTGAGTAA
AGGCGATACTGGGCATCGGCACACGTTTTATTCTCTTACGTTAGTACCTGCGC
TGGAGCGGCTTTCTTTACGCAGTAATAGCCGTATTTTCCAGCAGCAAAGCGTG
CCGGAAATTATCTCTATTTTACTGCAAGAGATGGGGATTGAAGATTACGCCTT
TGCCCTTAAGCGTGAGTGTGCCCAGCGTGAGTTTTGTGTGCAGTACCGTGAAA

CGGATTTGCAGTTTTTACACCGTATCGCCGCCGAAGAAGGCTTGGTATACAGT
CACCTACATGAAGCGCAGAAACACACGCTACTGTTTACCGATAGCTCGGACA
GCCAACCGAAGTTAGCCAAGCCAGTACCTTATAATGCGCTGGCGGGTGGCGA
GATCAATCTCCCTTATGTGGTCGATCTGCAATTCAAGACCACCGCACAAGTCA
GCCATACCGAGCTAAAGGATTACAGCTTTAAAAAGCCAGCTTACGGATTTAC
TCAGCGCACGCAGGGTAAAGATATCGCTTATCAGCAGCCAAATTATGAGCAT
TTTGATGCTCCGGGACGCTACAAAGACGATGCAAATGGCAAGGCCTTTAGCC
AAATTCGCTTAGAGTATTTACGTCGCGATGCTTTGCTTGCTGATGCGAAGAGT
GACGAACCTTTGCTGTTGGCGGGAGTGCGTTTTGACTTGCAAGATCACCTTGA
TCACGCAATGAATCGAGATTGGCTCGTGGTACAAGCCAATCACCAAGGGACG
CAGCCACAAGCGTTACAAGAAGAGGGTGGTTCAGGCGCTACCACTTACAGCA
ATCAGCTAAAACTTATTCCCGCTCACATCACTTGGCGAGCAAGGCCTTGTGCT
AAGCCGCAAGTGGATGGTCCCATGATCGCCACCGTGGTTGGGCCACAAGGTG
AAGAGATTTATTGCGATAACTTTGGTCGCGTGAAAGTGCATTTCCCGTGGGAT
CGCTACTCAAGTAGCAACGAAAAAAGCTCTTGCTGGGTGCGAGTGGCACAAG
AATGGGCTGGTAGCCAATACGGTAGTATGGCGATTCCGCGAGTTGGCCATGA
AGTGATTGTTTCGTTCCTGAATGGCGATCCGGATCAACCCATCATCACAGGTC
GTACGTATCATGCGACCAATACCGCGCCTTACGCCTTGCCTGACCACAAAAC
CAAAACCGTGCTGCGCACTGAGACTCACCAAGGGCAAGGCTACAACGAACT
GAGTTTTGAGGATCAAGCGGGCAGCGAACAGATTTTGCTGCATGCACAAAAA
GATTGGGATGCGTTGATTGAGCATGATCACACCGAAGTGATTCGCCACGATC
AACATCTCACTGTGGACAATGATCGCTTTACCCGCATTCAGCGCAACCAACA
TTTAACGGTTGAAGGTGAAGTCCGTAGCAAAATCGCACTCGATAGCAGCCAT
GAAGTCGGCGCGTCACTACAACACAAAGTCGGGCAGCGCATTGCTGTCGAGG
CGGGTAAAGAGATTTCACTCAAAAGCGGCGCAAAAATCGTGGTTGAAGCCGG
AGCGGAGTTAACCCTAAAAGCTGGGGGCAGTTTTGTAAAAGTCGATGCTGGT
GGCGTGCACTTAGTCGGTCCTGCCATTAACCTAAACGCGGGCGGCAGTGCAG
GCAGCGGTAGCGCTTATGGCGGGCAATTAGCGGCAGCACCGAGAATGTTAGC
GCAAGCTAAACCAGTAGCAGAATTGGTTCAGCCGGATATTGCGGCCTCAATG
CAATCAGGCGCGGCGCGTGTTATTGATGTGGCATCACTCCCCACAATGATGC
CGAGTTCAGCCAATAACACGGCAAATGATGAGCCTGTGGCCGAAGAGAAAA
CACCGGAGCGAATTTTAAAATCCGATCTCCTCAAACCATCCGATGAGTTAGA
GAAACTCGCCAAGCGACAAGCGAGTGCTTATCGTCAAGGTAATCACAGCGAT
GAGGTTAAGTTACTACAAGAAGCGCTCATAAAATTGGGGTTTGATCTCGGGA
AAGCAGGAGCTGATGGTGACTTTGGGAGCAAAACTAAAACCGCGATTGAGC
AGTTCCAGAAAAGCTATCAACCAAGCCACCAAACCCACCCGTCTTACAGCAT
TGGCGCTGTTGACGGTATTGTCGGCAAAGGTACACTACTGGCACTAGACGAA
GCCTTGATGGATGGGTGGGTGTATGAGAATAATATTTATCAAATATGGCCTCT
AGGTAAAACGTCAGAGAAATATGAGTCGGCAGGAAGAGGCCCAGGGGTTAT
ATCAACAGGAAATGGAGATTATGGTGGGGCTTCATATGGTTGCTATCAAATG
TCATCAAATCTTGGTGTGGTACAAAAATATATTCAGTCATCAAAATTCAAAG
AATTCTTTAGTGGATTAAATCCCGCTACGAAAGAGTTTAATGTTGTTTGGCAG
GATATAGCTTCAAGATATCCTCAGGAATTCAGAGAAGAACAACACCAATTTA
TTAAGAGAACTCATTATGATATACAAATAGGACATCTTAGAGGGAAAGGACT
TTTGTTTGAACATAATCGAGCTGCTGTACATGATTTGATCTGGTCTACTTCAG
TACAGTTTGGTGGGAGAACTAATTTGATTTTCAATGCATTGAATGGACAAAAT

ATGGAAAGTATGACTGATAAGGACATCATTATTCTGGTACAAGATTACAAGC
TTGTTAATACAGAAAGGCTTTTTAAGTCTTCGCCATCATGGTGGAGTGATTTA
AAAAAACGTGCGGTATCAGAGAAAAAAGCTTTACTTGAACTAGAAATTGACG
GTTTGGAGGTTGATATAAAATGA

C6706 large cluster immunity protein-encoding gene

ATGAATAACTTGCTTTCTGCGTATGTAACTATGCTGTTGATATTATTAAGTATT
AGTGGTGGCGCTATTGCTAGTGAAAACTGCAATGATACATCTGGAGTGCATC
AAAAAATTCTAGTGTGTATCCAAAATGAGATAGCTAAGTCAGAGACTCAGAT
TAGAAATAATATTTCTTCTAAGTCTATAGATTATGGCTTTCCTGATGATTTCTA
CAGCAAGCAAAGATTAGCTATTCATGAAAAGTGTATGCTTTATATAAATGTC
GGTGGCCAACGTGGTGAATTGCTAATGAACCAATGTGAACTTTCGATGCTAC
AAGGCCTTGATATTTACATTCAACAGTATATAGAGGATGTTGATAATAGTTAG

Amazonia, auxiliary cluster 2, effector-encoding gene

ATGTCTACTAAGTTTGCTTCAAGTCCGAGCTGTGATGGTAAAAAGTTTTTCAT
TGAAGTAACGGGTATACAACATGGCTCTGAACAAGACTTTGAGTTTTATGAC
TTAACTGATATGTCACAGCAGGCTGCTCTAGAAGCGAAAAAGAGTATTGATC
CTGAACTTGATGAAAGTACGGTATACAGTTGGGATTGGTGTGATGAGAGTGC
CAATCGAAATGTTTGGCTGAAAATTGAAGCGGAAGGTGATCCAATCAAATTA
CCGCTTTTTCAAAATGTAGCAGACATACCAAGAAAGAAAGACGAGCAAGATT
ATTTGGTACATGCAGTTTTACCTTTAACGCTTTTACCTACTTATCAATCATCGT
TAACTTATAAAGAGAGAATTGCTCCTGTTCGAGCAGGTTTTATTTATATCTTT
TACAACCATAAAGTATGGCGTGAAATTCAAATTAGCCCAGAAGATAGCGGTG
ATTACTCTCTTAAAGATGTCAATCTTTACCAATATCGTACTGGTAGAGATAAA
CCGTTTAAAGATGAAGCTCGAATCGCAACTGGTGTGGCACTAAAAGAGATTT
GGATACCGGCTAAAGAGAATAACAAGGGGGCACGGATCCATCTTGCCTTTTC
AGAAGTTCAGTGGAGTGCACAATACCTTAATTATCTTGAGGCGAATCAAAAC
GAGTTAGTTCAACGAGCTGTTGCATTTCATCAGTTGAATGTGGATAGCAATGT
CGATGTATTGAAAGCTAGCTTACTGCCGCCAATGCGAATAAGAGCGCCAGAG
TTGGAGTTGTTTCTTGCTGAGCCAAGTAATCTCAACCGTGATCTCTCTGGGGA
GTGGGTAACACGCACTTACCAAACCATTAAAGAGGAAATCTTATCCGCGAAT
GACGACGGTGATAAGGCGGTACAGGTTTTTAAATATGCTCAACCTCATCGAT
ATGAATATGCTATTAAGCAAGCCGCACTGATTGAGATTATCAACTCAGATCC
AAATCAAGCTAAGTTGTGGACTATTGGTGACTCGACAGATTTTTTAGAAGAT
GCAAAAAAGCGTCATCTGAGAGCAATTGTATTGGACGATCCTCTTTTTGACTT
AAGACATCACGCTTTTCTTACTCAATCTGCTGTTGGGTATTTACAGCAAGTCT
ATATCGATATGAGTCAACAAAAGTACTATCGCACTGCAGAGCTAGTGCAAAA
ATTGGTGGTTCCAGCCAAGTTCGGACAACAAGAAAATCCTTACTATCAATAT
CGTAATGCTATAGACAATTACTATGGAGGCGTTTTCCACAGAACATTACGTAC
TACTGAGCGACAGTTTTGTTGCCAAGATGTAAAGGTATTGCAAGAGAAACTC
CAATTGCAGGTGAACCAGAAGAGATTGGCTCATGTGCTGCGTGATATCAGTT
CAATGAACGATATTAACGCATCGGCTGCGCATGTGATTGTTGGCTATGCGCTG
AGTGCTCTGAGTGTGAATATTGATAAATTGGATCAAATGTCAAACCCTGAAA

GTGACGTTCGTTCACCTTATTTAGAAACCGCTAGACAAATCCTCTCTCCAAAA
GATAACCATCCTTTGCACCGAATATTATTTCCAGAAGAAGGTACGATAAATTT
GGAGCAAGCTAGCTATAGCGCACCACTGCCATTTAACTCGGGATCTGGATTT
GCAACGCTTGAATCGTTAGCTTTGTGGTCGAAAGAGGAGATGCTCATTCAAG
ATGAGCAGTTACAAGTGATGGATCTAGCTTTTATGTCGCCTTCCAATAGCAAC
CAGGAAAAGGCTTTTAACCTTGAACGGCGTATTGCCAGTATAATGAATGATA
TTTTAAAGGGTTACTTTGATACGTTGCTGAACCTAAGCCAAGATCTCATCAGT
GAAGCGAAAGTGATCCAATTTAACGCCGCTTATGCCCCTGTACTCGGTTTAAT
GAAAGCCACTAATTCCAAAATGTGGGGAGATATCACTTATATACCCGTCAGT
GGTGCTGAGCTGAAAGGTACGGTGGTTGGCGTGCATGGTCATGGCTTGAGTT
ATGGATTGAGCGCCTCTGATCGCGAGTTTGTTGACTCGAAGAAGAAGTCAGC
GTTTGGACGCTTGTATGATCGTTCGGGCAAATTGGTGGCATCGACCAATAAA
AATGCTTTTAGCTCCAGTGATCTGATTTCAGAAAGTCGAGGTTCAGTCGGGGC
TAAATCACCACTTAAAGTGGTGGTTGTATCTCAAGAAAGCCAAATGGCAGCT
GCGTTCAACCAAGCGAATACGCAAAGAGCGTTACGCGATATGAATCGTACGG
ATCTCAACGTCAGTAATGCTTATGAAAAATTCAGAATTCCTTACTTTATTGCG
GTAGTGGAATTGATTAACCTTAAACAGAGTCAGCATCATTTTGAGAAAGCGC
TTCAGCGTACAGATACTCTTTACTCTTCTGCTAACGCGCTAAGTGCCGCGGCG
GATCTTACTATTGCGCTCATTCATGCTTCAAATCTGTATACACAGAACGCCTC
TCGTTTAGCGACAGCCTCAGGTAAAGCCGCGGTTGTTATGCCTGACTTCCTTG
TGAAAAAAATGACGTTTAGAAATGGTAAGGTGCGTTTGGTTCCCCACATCAG
CCGATTAGGATTAGCAAGTATTGGTGCTGGTTTTTTAACTGCGGGTATTGCTG
GTTGGGATGCGATGCGCCGTTGGCAAGAAAATGATCTTGATGCCAGCGTTGC
GATGGGTATGGTGGCGATAGGAACGTTAACCACTACGGTTGCAACGGGTTTT
TTTACCACCAGTGCTCCGGTGTTATTTGGTATGGGACCGGTGGCATGGTTAGG
GATTGGGCTGGCCGTTGCAGGTTTCGCGCTCTACATGTTTTGGAAGGATACCC
CAATGGAGGCGTGGCTGAAAAATGGCCCGTTTGGTCAGTCACCCTCGGCGAC
TTACGCACATTTGCAAGACCCGACGACGGCGTTTGAGCGGTTTATTGGGTTGA
TTTTTACTCTATCGGTCAACGCGTACCGCTTAGGTGCACAGACCGAATTCCCC
GAGGTGTTTACTCAGCAGATGCAAGCATTAGGTGCTACCCATGTTATCCATGT
CAATACCAATTTAGCGGCGTTACTCAATTCTCAATCGGTCAGAGTTGAGTTTT
ATGCTCGGCAAGCGATTGAAAAAAAGACCATAACAAGTAGCCGAACGGGCT
CAAGGGAAAGTAGTGAACTGATTAACCTAAGCTCGCACAATGTTACGGTTAT
CCATCAAGAGCAGCGTAATGAAGGTACTGCCTATTTTGTGAAATATGATTTA
ATCGTCCCCGAATACAGTACCGATTTTAGCCTACTGATGATGCGGTCTTATCA
GTATCGCTATCAGCCTGTTTTTGTACTCAGAACCAAGCTGCACGTTGAGCAGG
CTAGCTTTCCGACATTAGCATTAGAAGAGCGTGATGATTCTCGTATTGTGAAT
ATGGTTCCAACGTTTCGTTTAGATAACAACAGTGATCAAGATTGGGCGCAAA
ATACCGTATTAGCGTAA

Auxiliary cluster 2, Amazonia, immunity protein-encoding gene

ATGGCTTATAACTCAGAGCAATTAACCCTCCAACCGCCTAAAGAGGCGCAAG
GTTTTTTACGGAGTGTCTTAACCAAAGGGCAGATATTCCCTTTTACTAAAGCG
AATGAGACCATTGCTCTTTCTTTGGCTGGGGAAATTACACCTTCGTATATCAA
AGAGACTGAAAAGGAAGACAAGCTCTCTTATTGGGATGAAAACTCTTTAGAG

AGTGAAACTTTTTCGAATTGGACTCAGTTGTACTTATTTATTGTAGGCTTAGC
TAAAGTAGCATTGTTTTTTTTCTTACCGCTTCTCTACCTGATCATGATTTCTGC
CATGTTATTTGGACCGTATGGATGGAAGGATTGGGCTGGCGATTTTGGCACTG
TTACTTTGTATACAACACTTCCTTGTTTACTTATTTATGGCCATTTTAAGTTGG
TCAGTGCTGGGCATCTTTTTCTGGCGCCCTTTTTAAAAAGTAAACGAGTGTAT
AGCCTCAATCGCCAAACTGGCATGGTGACGCTGTTTAAAAAAGGTAATAAAG
AGCGTTTTAGTCATCCATTTATCGAGTTTGACTGTGTGCTGATGTCCGCTCCTA
GCCCACAAGGGAATTTAAATTACAACCTGATGTTGGTGCATCGATACCATGA
TTATTCCGTTGGGGTGCCGATCGGTAATCTAATTGGCAGTAATGAGATGGTGG
CAGAGTATTATCGGTTGTGGAATATGATCCAACGTTACATGGATATCAGCCA
ACCCATGCCCGATATTTTAGTGCTTGAGCCTGCACGTGAAAGAGACCCAACC
ACTGCCGCCTATGATAAGCAGACTGGGCGCAATCCGCGATACTGGCGTGATA
TGAGTGATGAAGAATATCAGCAAACCCTGAAAAAAATCGCAGATAAGCAAA
AAAATCAGCCAGATAGTGGCCCAGTGTTGAATGTTTTTTCACTATCTTAG

Auxiliary cluster 2, MZO-2, effector-encoding gene

ATGAGTAATCCCAATCAAGCTGCGAAAACAGGACAGACCAATGATGCGCAA
AATCCCGCCAGTGCATGTCCTTTTAAACAGCCGTTAATAGGGATAATTCCTGT
TCGTTACGCCTTTGATGTTTATGATGATCAAGGTCAAGCATTACATCCTTTAC
CCAAAGCGGATCGACAGTGGAAAGGTCAATTCTCTATCAAGCAGCGCAGTTA
CACTTTAAGGCAACTGCGTGATGGCTGGCTCTATGTCTATGATGAAACAGCA
AAAACACTGCACGAATATGAAGTCGTTGGCTGCAAGTTGACCAAAATTGATT
GGTCTGATGATGAGGCAAACAAACCCACTCACGAGCGCGGTTCGAAAGGTGA
AAGCAAAAGCTGTTTACTCTATCCCGCCCAGCACACGCTCTCTATCGGTTATG
CGCATCAACGATGGACGTGGCGAGTGTGTGAGCATATGCGCTCCAATACCAG
CAGCCGTCATGCTGTGATGCGTAAAGTGAGCCTTAAACAGTTTGAGTCCAAT
GGTACTCACCCTCATGCACATTTTGCTCAGTATCTTGAAGACTATGTGGCGGA
CATCGGGACGCCAGCTGAGCAGGATATTTTTAAAGATACCTGTACCCCTTCAT
TACCGGTAGAGAAAAGTGAAGAAGCGGTGAAAGGTACGGAATTCAAATTCG
TTGCCGACAAAGCGGTGGTAAGCAGCAGTGATTATTTGCAAGATTTGCCGGA
GCAAAACTGCGGTTTATTTGTGGCGCTGAATGACCCTTTAGCGGATGTGTCAG
ATTTATTTGTTACTTTTACTACCCAAGTGGCAAAGCGAACTAAAGCGATTGGT
GATGAAACGCAACAGCATAAAATGCAGATGGCAGAGCTGACACGCACATTA
GGCCGAATTCGTCTTGAAGAAAAAGAAATTCCCGATTTTGTAAAACAAGATC
CGATCCGTATTTTAGAGCTAGAAAGAGCCATTACTGAATATTGTGCTACGGC
AAAATTGGCCGAAATCGAATCTCATCATTTAGCCAGTGAAGGGCATTCCCCA
TCGGGAAACTATGCCTTGATGCAGCAGCAAGCGGAGCAGAAACTGGCGGAA
TTGAAAACACTCTATCGCTTTGAACCCACCAGTGCTCAGATGCGCAAGTGGC
GCAAAAAAGACAACAGCTTTATTGATGAAGTGCGTTGGGCGGATTTAGATAA
CTTTTTGGTGGAGCATTACACCGAGCTCAAAGGATTGGATGAACAGATTAAA
CAACACTATGCGCAGTTTATGTCTGCCTTCAACCAGCTCGGCCTTGACCCGCT
TCTGTTTGGTATGGATAACCAAGATGAGGTTCAGCAAGCGTATTTGTTAGCAT
TAACTAGCCAATTTTTAGTGGTGGTTACGCAAGTCAATCATGATGAAAAATC
GCTTGAGATCCTAAAAAAAGATCTCAGTTTTGATTCACCGAAAAACCTCATG
GCCCTCGCGTCAACAGGCTTTTCCCTGCAGGCAAATCAAGCCATTAACAACC

ATATCCAAGGGTTTAGCACCGCTTTTCTCTCTACAAGTAGTCCTAGTGATATG
GTGGCGTTTGCGACCGCAATTGCTAACTGGGATACGTTCACGGGGGATGAAC
GTATTCAGGAGAAAGCATGGTTTAAGAGATGGATTGAACCCGTCCAATCGAG
TTTTGGTGCACTGAAAAAAGCGGTTGCCAATCAAGCCAAAGAGAGTTGGCAA
GCGGTGATGGAGTTGCTGTTTCCTTATCAGAATCAGCCCAAAGGAGGGACGC
CCAGTTTGTTGGCTAACCTTCGTTTACTGTTGGTGGAAAGCCTAGTTCGAGAA
GAGGCGGTGCTGCAACATAACCCCAAGTATGCGGCTGAGCTTAAGCAATTTG
AGACTAAGCTGAATGCCATTTTGCAAGAGATGAATGATGCTCTGGAACTCAA
ACCCGGTAATGTGAGCCCCAAAAACCATCAAATTGCGACCGCGCAATCCGCA
CAACGTAAGCTAGGGCAATTACTCAGTAGTGAACTGCCGATGATGCTGACGC
TCAAAAACCAAGCGGTGATGAACACTTTTCAGCAGTCAGTAAATGAGAAACT
GAGCGCGCTATCGAAAAACGTCAAAACCAGTAGCGCGTCGGTGAGTCAGAA
ACTGGGTGGATTAGGAGGCTTGCTGTTTGCGCTGAACCTATGGAATACCATG
ACGGTTTTGGAGAATATCCGCTATAAGGTCGCACAATACCCCAGTTGGAATC
CTTTTAAGAACCCGGCGTTGGGCGAAGCCATTTATGCCACTGGCTATACCATT
GTGGCCGCTGGTGCCATCAGTGCGGGGCGCGCTTGGGTGACGATTGTTAAAT
ATGAGCTTTTGGATAAATCTCTAAAAGAAGCATTAGGTGAGGCATCCAGTCT
TAAAGCCAAAGATGCCCTAAAAACTTTTTCTAAATCCATCGCTTTGGTGGCTA
CCGTGGGCATGATTGCCAGCGCATTAGAAACATGGGAAAGCTGGGGTAAGTT
TAATGACAGCAGTAAAACCGATCTTGAACGCTTCGGTTATTTATTAAAAGCA
GGGGCAACGGGAGCACAGGCTCTGATTTTTGCTATTCAGTTAGGTGTATACG
GAGTAAGTCGATTCATAGGTTTTGGAACAATCGCAGCCATCAGTGCTGGTTG
GATGGCCGCCGGCTTTGCTGTGATTGGTATTGTCTATTTGATCGGGGTGATTT
TGACCAATGTGTTTAAACGCTCAGAGTTAGAAATCTGGTTATCAAAATCGAT
GTGGGGGAAAGAGAGTGCTCATTGGTCGGTGGGCAAAGAATTGACCGAGTTA
GAGCGCCTTCTCCATCGTCCGAGTCTTCGGTTGAGTCAAGTTACGCAACGTAA
AGCAGCGCAATGGATGGATTCAGGTTCTTTACAATGGCAGTTGGAACTCACG
TTACCCGATTATCTTAAGGGGCAGACGATAGGCTTACAAATCACCCGTTTGCC
AGCGCAGCCGGCATATTATCAACGGCAAAGGGAGGCGGTAACACCAATACT
CATCAATGAACAACAGGGAAAATGGAGTATTGAAGATAATCAGCCAGTGTAT
CGCATTACACTGGGTGGCAGTGAGAAGGATACCGTTGGAGTTTGCGTTGCAC
TGCCCCTGCACTGGGTCAAGGAACAGAGTCTCAAATTTTGCGCCCGCGGAAC
TCGCGTGGGGGAATTAGATTTACAATCTGCAGAGGCGAATGATATTGCCACA
AGAAATCTCGTTGTAGGAAAGGGTTAA

Auxiliary cluster 2, MZO-2, immunity protein-encoding gene

ATGTTAATTGATAAAGATGAGTTAGAGCAGTTAAAGGCAAATTTACACAGCA
GTGAGGTTTTGTATCAGTGGGAATGTGTGGCATTTTCAGAGCGACGAGGTGA
TATCGCTCGTTGGATTATTGCCGTCATTGCAGGCTTTTTGCTTCCCTCACTATT
TCTTGTGTTCGGGGATATTGCTTTTAACAGCACTGAGTTTTGGGGGTTTATCA
GCTTTGGCTTGGCGGGCATGTCGGCGGGAAGATATCTTTTGATGCCCGACCAC
CGTTATTGCTACAGTCTAACGCAAGCGGGGATTTATTATACCGACCAAGAAG
TGATCCCTGATGCGGCCTACACTTTTGTGCGTGGTTTTGCTTGGGTGGGGATT
GCGGTCTGTTTGCTAGCCTTAGCCGTCGTTGGCCCGTTAGCGTTTGTTGGTGC
AGGCGGTTTTGCCCTGCTGGCCTTTGGTTTGACTAACTTCCATCCCACCGTAC

ATCAGAAAGATGTTTACTTTGCTGACCAATTGATCCTTTTTGATCCTATCAAA
GAAAAGATGGTGAAATTGAATTCCGAAAGCAGAAGACATCCTCGATTTAGTA
GAACACTTTTTTTTAGTTCATTTGATGAAAAAACACGTTTTATTGAGCTCGTG
AAGAGTATTCATATTAATGTCGACTATCTTCCGTTACAGCGAGTTAACGATCA
GTACAAGCATCCAATCTTTAATCAAGAATTAAAAGAGGAATAG

Auxiliary cluster 2, C6706, effector-encoding gene

ATGAGTAATCCCAATCAAGCTGCGAAAACAGGACAGACCAATGATGCGCAA
AATCCCGCCAGTGCATGTCCTTTTAAACAGCCGTTAATAGGGATAATTCCTGT
TCGTTACGCCTTTGATGTTTATGATGATCAAGGTCAAGCATTACATCCTTTAC
CCAAAGCGGATCGACAGTGGAAAGGTCAATTCTCTATCAAGCAGCGCAGTTA
CACTTTAAGGCAACTGCGTGATGGCTGGCTCTATGTCTATGATGAAACAGCA
AAAACACTGCACGAATATGAAGTCGTTGGCTGCAAGTTGACCAAAATTGATT
GGTCTGATGATGAGGCAAACAAACCCACTCACGAGCGCGGTTCGAAAGGTGA
AAGCAAAAGCTGTTTACTCTATCCCGCCCAGCACACGCTCTCTATCGGTTATG
CGCATCAACGATGGACGTGGCGAGTGTGTGAGCATATGCGCTCCAATACCAG
CAGCCGTCATGCTGTGATGCGTAAAGTGAGCCTTAAACAGTTTGAGTCCAAT
GGTACTCACCCTCATGCACATTTTGCTCAGTATCTTGAAGACTATGTGGCGGA
CATCGGGACGCCAGCTGAGCAGGATATTTTTAAAGATACCTGTACCCCTTCAT
TACCGGTAGAGAAAAGTGAAGAAGCGGTGAAAGGTACGGAATTCAAATTCG
TTGCCGACAAAGCGGTGGTAAGCAGCAGTGATTATTTGCAAGATTTGCCGGA
GCAAAACTGCGGTTTATTTGTGGCGCTGAATGACCCTTTAGCGGATGTGTCAG
ATTTATTTGTTACTTTTACTACCCAAGTGGCAAAGCGAACTAAAGCGATTGGT
GATGAAACGCAACAGCATAAAATGCAGATGGCAGAGCTGACACGCACATTA
GGCCGAATTCGTCTTGAAGAAAAAGAAATTCCCGATTTTGTAAAACAAGATC
CGATCCGTATTTTAGAGCTAGAAAGAGCCATTACTGAATATTGTGCTACGGC
AAAATTGGCCGAAATCGAATCTCATCATTTAGCCAGTGAAGGGCATTCCCCA
TCGGGAAACTATGCCTTGATGCAGCAGCAAGCGGAGCAGAAACTGGCGGAA
TTGAAAACACTCTATCGCTTTGAACCCACCAGTGCTCAGATGCGCAAGTGGC
GCAAAAAAGACAACAGCTTTATTGATGAAGTGCGTTGGGCGGATTTAGATAA
CTTTTTGGTGGAGCATTACACCGAGCTCAAAGGATTGGATGAACAGATTAAA
CAACACTATGCGCAGTTTATGTCTGCCTTCAACCAGCTCGGCCTTGACCCGCT
TCTGTTTGGTATGGATAACCAAGATGAGGTTCAGCAAGCGTATTTGTTAGCAT
TAACTAGCCAATTTTTAGTGGTGGTTACGCAAGTCAATCATGATGAAAAATC
GCTTGAGATCCTAAAAAAAGATCTCAGTTTTGATTCACCGAAAAACCTCATG
GCCCTCGCGTCAACAGGCTTTTCCCTGCAGGCAAATCAAGCCATTAACAACC
ATATCCAAGGGTTTAGCACCGCTTTTCTCTCTACAAGTAATCCTAGTGATATG
GTGGCGTTTGCGACCGCAATCGCTAACTGGGATACGTTCACGGGGGATGAAC
GTATTCAGGAGAAAGCATGGTTTAAGAGATGGATTGAACCCGTTCAATCGAG
TTTTGGTGCACTGCAAAAAGCGGTCGCCAATCAAGCCAAAGAGAGTTGGCAA
GCGGTGATGGAGTTGCTGTTTCCTTATCAGAATCAGCCCAAAGGAGGGACGC
CCAGTTTGTTGGCTAACCTTCGTTTACTGTTGGTGGAAAGCCTAGTTCGAGAA
GAGGCGGTGCTGCAACATAACCCCAAGTATGCGGCTGAGCTTAAGCAATTTG
AGACTAAGCTGAATGCCATTTTGCAAGAGATGAATGATGCTCTGGAACTCAA
ACCCGGTAATGTGAGCCCGAAAAACCATCAAATTGCGACCGCGCAATCCGCA

CAACGTAAGCTAGGGCAATTACTCAGTAGTGAGCTGCCGATGATGCTGACGC
TCAAAAACCAAGCTGCGATGAACACTTTTCAGCAGTCGGTAAATGAGAAACT
GAGCGCGCTATCGAAAAACGTCAAAACCAGTAGCGCGTCGGTGAGTCAGAA
ACTGGGTGGATTAGGGGGCTTGCTGTTTGCGTTGAACCTATGGAATACCATG
ACGGTTTTGGAGAATATCCGCTATAAGGTCGCACAATACCCCAGTTGGAATC
CTTTTAAGAACCCGGCGTTGGGCGAAGCCATTTATGCCACTGGCAACACTATT
GTAGTCGCCGGTGCCATTAGTGCTGGGCGAGCTTGGGTGACGATTGCTGAGC
AAGGGTTGTTAGATCGTACACTGAAAAACGCCTTAAATACCACGAAAGTTTT
GGGCACAAAAGATGCCCTAAAAACCTTTGCTAAATCGATCGCCTTGGTGGCT
ACCGTGGGCATGATTGCCAGCGCATTAGAAACATGGGAAAGCTGGGGTAAGT
TTAATGACAGCAGTAAAACCGATCTTGAACGCTTCGGTTATTTGTTAAAAGCT
GGAGCAACGGGAGCACAGGGAATTATTTTTTATATTCAGTTTTTCACACTCTT
AGGTAGTGGAATTGGTGGACCCTCCATAGCTGCGATTTCTGCAGGTTGGATG
CTCGCCGGTTTTGCTGTGATTGGTATTGTCTATTTGATCGGGGTGATTTTGACC
AATGTGTTTAAACGCTCAGAGTTAGAAATTTGGTTATCAAAATCGACGTGGG
GGAAAGAGAGTGCTCATTGGCCAGTGGGCAAAGAATTGACCGAGTTAGAGC
ACCTTCTCCATCGTCCGAGTCTTCGGTTGAGTCAAGTTACGCAACGTAAAGCA
GCGCAATGGATGGATTCAGGTTCTTTACAATGGCAGTTGGAACTCACGTTACC
CGATTATCTTAAGGGGCAGACGATAGGCTTACAAATCACCCGTTTGCCAGCG
CAGCCGGCATATTATCAACCGCAAAGGGAGGCGGTAACACCAATACTCATCA
ATGAACAACAGGGCAAATGGAGTATTGAAGATAATCAGCCAGTGTATCGTAT
TACACTGGGTGGCAGTGAGAAGGATACTGTCGGAGTTTGCGTTGCACTGCCC
CTGCGTTGGGGCAAGGAGCTGAGCCTCAAATTCTATGCCAGTGGAACTCGCG
CGGGGGAATTGGATTTACAATCTGCAGAGGCGAATGATATTGCCACAAGAAA
TCTCGTTGTAGGAAAAGGTTAA

Auxiliary cluster 2, C6706, immunity protein-encoding gene

ATGTTAATTGATAAAAATGAGTTAGAGCAGTTAAAGGTAAAATTACACAGCA
GTGAGGTTATTTATCAATGGGATAGTGTGGCTTACGGAGAGAGAAGATCTGA
AATATTTCGCGTTTTTGGGGCAATCTCTGCTGGTATTGTTCCCCTTTGGCCATT
TATTTTTTTTGCAGATATTCAATTTAATAGTAAAGAGTTTTGGGGATTTATCTG
CTTTAGTCTGGCGGGCATGGCTGCTGCCAGATATCTTTTTATGCCCGACCACC
GTTATTGCTACAGTTTAACGCAAGCGGGGATTTATTATACGGATCAAGAAGT
GATCCCAGATGCGGCCTACACTTTTGTGCGTGGTTTTGCTTGGGTGGGGATTG
CGGTCTGTTTGTTAGCCTTAGCCGTCGTTGGCCCGTTAGCGTTTGTCGGTGCG
GGCGGTTTTGCCCTGCTGGCCTTTGGTTTGACCAACTTCCACCCCACCGTACA
TAAGAAAGAGGTTTACTTTGCTGACCAATTGATTGTTTTTGATCCCATCAAAG
AAAAGATGGTGGATTTAAATACAGATAGCACAGATGAACCATGGTTTGACAG
ACGGCTATTTTTTAGTTCACTTGATGAGAAGACGCATTTTATTGAGCTCGTGA
AGAGTATTCATAATAATGTCGACTATCTTCCGTTACAGCGAGTTAACGATCAG
TACAAGCATCCAATCTTTAATCAAGAATTAAAAGAGGAATAG

Auxiliary cluster 1, Amazonia, effector protein encoding gene

ATGAGTAGCACAGGTGCCCATCCACATTGTCAACCTTGTGAGAATCTCAAAC
ATTGGATTGAAATTATTGTGCGGGATGAACATAACCAGCCCTTTGAAGGTGT
GAGTGGTGTTTTGATTGACGCAATGAAAAATAAGCATCCAATTGAACTTAAT
GCTTCACCTATCCTGATTGAAAATCTTGCGCCCGGTCCTGTTGAAATTGAGTT
GGATTACGATCAGTGGTTAAAAGCTGCGCAAGATAAAAGTCACCCAAGAAAT
GAAAAGAAAGCTAAACTGGTTGAAGAGTTTTCTAACAGTTACAGTGCCCACA
AGAGTGGCCCGCTGGTTTTTCAAGAAATTACGACGGGTGATTTGACCAAATT
ACCGAAAGAGATCGTATTGCCAACTAACCATCAAAAAGGCAAAGCCGGAAC
GTTAAAACTCTTTACAGATAAAACCTATATTCTTCAAGTCAGGGCATACAAGT
TTATTACTCTTCGAGTCGGAATGTTCTTTGATGGTACCGCTAACAATACCTAC
AGCGCCCAATGGGGAAAACAACAATTAGAAAATTATTATCGTAAATGGAAA
GCAAAGTACGATGCTGAATGCGAGATTAATTCAAAAAATAGTAATGGAACAA
AAAAAGAAGTTCCAATTACGGCACTTTCTAATGATTGTTTTACCTATCCGAAA
AAAGAAAATTTTATTCTCTCCTTATTTAAAAATGATGAAGGAGAAATGGAAA
CGGTTGCGGGTAGTGCAAGTAATGAATTAACCAATGTTCAAAAACTGTTCGA
TTTATATTCTCAAGACAAATTTTTTAAAGAAAAAAATATGTTTAGCCATGCTG
AGTATATCACTGGGATTGGTACGGGAAATAGCACTGCGATTGCTCCTGCAGA
TGAGTCGATCGTTGTCGGGCAAGGCTTAGGAATTGGTAAGTATGGAGTAACG
GCTAAAGTTACAACGGGAATTGAAGCTCTTTCTAAGAATATGGATAAAGTAG
CGACTATAGTCAAAGATGAACTCGGAATAAAAGCAGATGGAATCGAAAAGC
TTCAACTCGATGTGTTTGGGTTTAGTCGAGGAGCTGCCGCAGCAAGGCATTTT
GTGAATGTGGTGCTGGATGGTGAAAAGGGTGAATTTTCGACTACGTTTTCAA
AAGCATGTCAGGAAGCTAAATTTCCGCTAGTGTATGGTTTCGATTGGAATGA
ATCGAATGAATTAAAAGCCAATTGTGAAATTACCTTTGCTGGTTTGTTTGATA
CAGTGGCCTCAGTAGTCAATATTTTTTCTAAAAATTCCCCTTTGGGCCTAGAT
CTCAATACACACACGGATAATGGTGATGTAAGACTATGGATTGACCCTAGAC
GAGTCAGGCGTGCTGTACATTTAACTGCGGATCCGACCATTGAATGCCGAGA
TAACTTTAGTTTGAACCATCTTAATTCTACGGATGAAGAACATTTTCATGAAT
TTGTTTTGCCAGGCGCTCACTCAGATATTGGGGGCGGCTATCACTCTAGGCTC
AGTTTTGACAATCCAGACTATTTGTTACCTGTTTTAGAGAAGAAGCTTGTTAA
GCGTGTGAGTCGTACGTTTTCCGAACGCTGGGACGAAGAGAAAACTAAGCAG
TATGTGTTAAATGAGCTAGAAAAATATAAAGTACGAGATAGGCTTACGGGGT
GGAAAGAAGAGGATTACGTGATTGAACCGTTGGACATTCGCCAAGAAGGTA
AGAACGATGGTGGGCGAGTGACAGGCAAACTTTATATTCAACGTCAAGTAGA
AGGCGACCTCTCACGTTTATACTTACGTTTAATGTACGGTTTGGCAGAGTTTC
ATGGAGTTCCGATTTCAGATAATAATGCAAAATTATGGCAAGACCCTGAAAG
AGTAGATTATAACGTTGAGGATTATGGCGGTTTATTTACCGATCTTAATCAAA
AAATCTTAGAACTTGCTAAACATGGGGAATATTCAGCGTTGCAACAGAAACT
CTCAATCCCAGAACTAAAGGCAAGTTTCATGGAATTGAATCTCTTTCACCATT
CATCGGGGGATGATATTGGGATGTCACCACTCTGGGATGAAAGAGTTGGTTG
CTATAAAAGAGCAAGTTATTTTTGTGAGGAAGGCAAATGA

Auxiliary cluster 1, Amazonia, immunity protein-encoding gene

ATGAAAAAGACATTGGTAATTCTATCCCTATTTTTAGTAAATGCTTGCACGTC
TAATGCTGCATCAGTTCCTGGTGAGCCTTGGTATGTCGGTGTTACTATGCCTT

CTTTTTATCCAGTACGAGTTACACAAGTTTATGGGGTAAATAATAAAGAGGA
TTGGACAGTCTTAATTCACAACTTTATGACATTTGTGAGTGATAGCGATGTTG
AGTATATACGTAATAGATTTCCAGACTATGACGGATTTGGTCTTCCACATGCA
TTAGGTATTACAAGCAGAAACCAAATTGGTACCGGAACCAATCACTTGCCTG
ACACACTCTATGTATATTGGGTTTCTTTATTCGACACTAAGTTTTATGTCACTA
AATATGATATTCCTGACAACGTAAAACAGTTGAGTGCGACAGAGGTCGGTTA
TACTCGAAGCGATGGTGAGTTCTTTGAATCTTGTTATCGTACTCGATTTGTCTT
TGGTCTTCTGCCTAATGGGCAAGCTAAGGTGTGGTTGAGTGGCTGTGGTGAG
ACTATTTATCTTACGGAGTTAGCTCCTGACGAAATACTTGATCGTGACAGTAA
TGGCGTTAAATTTGATTCTTATCGGAAATCTTCCTCTTTTGCTGATGTTCAGCA
ACGCGCTAAGGATGCAGGTGTTGAGCTAGAGCCTATCCCTTGGGATAAAGTG
AATAAAGTCTATTCCATCAATGAGATTAAAACGCTTGGTGAGTAA

Auxiliary cluster 1, C6706, effector protein-encoding gene

ATGGATTCATTTAATTATTGCGTGCAGTGTAACCCTGAGGAAAACTGGTTAGA
GCTAGAGTTTCGGAGTGAAAATGATGAACCGATAGATGGTTTGCTAGTTACG
ATTACGAACCAAAGTGCCCCCTCTAACACATACACGCAAACCACAAGTTCGG
GTAAGGTGTTGTTTGGCAAGATTGCGGCTGGGGAATGGCGTGCTTCTGTAAG
CCAAGCTTCTCTGTTAACTGAAGTCGAAAAATATGCCAGCCGTAAAGAGGGA
CAAGAGTCGCCAGTGAAAAAACGAGCTGCGGCAGAACTTGACGCCGCAGAT
AAAGATACGAAGCAGTACCGCTTTACCACTATTGGTGATTTCTGGGATGAAG
CACCAAAAGATGAGTTTTTACAAAAGCAGCATAAAGGGATAGATGTCAATGC
TTCTGCTGAGAAAGCAGGGTTTAGACTCAGTCATAACCAGACCTATGTTTTTG
AAATCAAAGCACTACGTAGCTATATGCCAGTGATTATTGATACCGATGAATTT
AACTTGGTGAATAGTTATACCTTTGCGTTGTTGTCGAAATTAGCTTATGCGAC
TAACGATTTTAATAGGGATGATGGTAAGACAATAGATAATCAGGGAGCAATT
AGCACGGTTATTTCTCAATTGAAGCGTAAAGAAAGACCGACTTATTCTGGAG
ATTTACAAGCAAAGTGGTTATTAGAGGAAATTCCATATAGCAAAGCGTTAAG
TGCCCAATATTATGCTGAGGATGATGTCGGTAGTGAAGGTTACATCATCTTTA
ACGATGAATTGGCGATTATAGGTGTGAGAGGTACAGAGCCTTATTTTCAGAG
TAAGAAACCACCTGTTGATAACACTAAATTTAAAATTATAAAAGCGGCTTCA
GGTATGGCGGCTGTGATTGCGGATAAGATAGAAAGTGCTACCGATTCACCAG
GAATGAAAGACCTTATTATAACAGACCTTGATGCTGCGCAAATCGCCCCCGA
AGAGTTTGGTGGAACATATGTTCATCGGGGGTTTTACCAATATACGATGGCTT
TGTTATCTTTGATGGAGAAAGATCTAGGGCTACATAAAATAAAAAAATTCTA
TTGTTGTGGCCATAGCTTAGGGGGCGCTGGTGCATTACTTATAAGTGCATTGA
TTAAAGATAGCTATCATCCTCCAGTATTGCGTTTATATACATATGGAATGCCA
CGAGTGGGAACTCGCAGTTTTGTAGAGCGCTATCAAAATATTTTGCATTACCG
TCATGTTAATAATCACGATTTGGTGCCACAGATACCAACAGTATGGATGAAC
ACAGATGTATCTGAAGGATTCCATGTACTTGATGTATTTAAAAGCCGTGTTGA
TCTGATGAGAAAAATGCTCACTGATGATGATGATGACAATTATCAACACCAT
GGTCATTTGTCTCAGTTACTAACTTATAATTCAAATAACCAAGTCTTACTGAC
ACCCAAACAAACTCAAGTCACTATGCTAGACCTAGCAAATTTAGCCACTAAT
GATTCTGTTGCTATGGTAGATGGTTTATCAGATGCCTCAATAGTAGAGCATGG
AATGGAGCAGTATATACCCAATTTATTTGAGCAGTTAACAGCACTAAGTGAT

GAAAGTTTAATGGTGCATTATCAACGAGCGATTTCAGCCTTAGAGCAGGAAA
TTGCTACGTTGCAACAAAGTTATCTTACAGTTAAACAAGCTTGGATTGAATCT
ATAGGAAATGGCACGCCGACTATGAATATCGGACGTTTGATGAGTGAAATGC
ATTCTATAAATAAATTAATAGAGAACCGTAATAAAATTAGAGGTGAGCTAAG
GCAGATAGTCTCTGATCCTCAGCGTATGCCTGCAACTAAATTTTTAATTTCAC
AGCAAACCTTACCAGATGAAATCAAGGTGCAAATAAGATGA

Auxiliary cluster 1, C6706, immunity protein-encoding gene

ATGAAGTTATTGAATAATCTTGCAATAAAAAAGCTAATCTCAGTTGTTGTACT
ATTCATTGCATTATTTTTTCTTGGGTATGACTATTACCAGACTTCTCAACCTAA
TATTTGGGGTGATGAACCTGATGAGTCCTATATTACTATATCAGGCAAAAAG
CCTATAGATGCTTATTTAGAAGTGTGGACTCATTTTTGGGTGACAGGAGATGA
GTGTGAAGCTTATTCCTACGATTTATTTGGTCAAAAAGCCCATCAAGGAGGA
AAAATAAGCCAGAAGATTACGCATGATTTTGCTAAAGATGGTTCAAATTTTG
AATTTAGAATTCCTTACCAAACGTATAAGAACTCACAAAATTGTATAGTGGA
ATTAAGAGACTTTTCCATACAAGCTCATAATGATTTTGATACCGTAGGTTTTG
CTCAGTTGCGATTTTCTCCAGCAGGCAGAGAATATTACAATCGAGAGGTTGA
TCTAAACTCATTGATTACAGCAGATAATTGCAATAGTGATATATTTAAAAGTA
TAAGAAAAGAGTGGGCAGGAGCCATTGGTTGTCATTTTTATGTCGATGGAAA
GAAAAAGACTAAAGATGAAGAATTCAACGCTTACACCATCTACTACGATTTT
TCGAAATTCAATAATGACACCGTGATTCATTACGATATTTTAGCCGGTGAAAA
CTATCGGTCAGAGCCTTTGTCCTCAGCTCAAAAGACAGCAGTGGTATCTGATG
ATAATTAA

Auxiliary cluster 1, NIH41, effector-encoding gene

ATGGATTCATTTAATTATTGCGTGCAGTGTAACCCTGAGGAAAACTGGTTAGA
GCTAGAGTTTCGGAGTGAAAATGATGAACCGATAGATGGTTTGCTAGTTACG
ATTACGAACCAAAGTGCCCCCTCTAACACATACACGCAAACCACAAGTTCGG
GTAAGGTGTTGTTTGGCAAGATTGCGGCTGGGGAATGGCGTGCTTCTGTAAG
CCAAGCTTCTCTGTTAACTGAAGTCGAAAAATATGCCAGCCGTAAAGAGGGA
CAAGAGTCGCCAGTGAAAAAACGAGCTGCGGCAGAACTTGACGCCGCAGAT
AAAGATACGAAGCAGTACCGCTTTACCACTATTGGTGATTTCTGGGATGAAG
CACCAAAAGATGAGTTTTTACAAAAGCAGCATAAAGGGATAGATGTCAATGC
TTCTGCTGAGAAAGCAGGGTTTAGACTCAGTCATAACCAGACCTATGTTTTTG
AAATCAAAGCACTACGTAGCTATATGCCAGTGATTATTGATACCGATGAATTT
AACTTGGTGAATAGTTATACCTTTGCGTTGTTGTCGAAATTAGCTTATGCGAC
TAACGATTTTAATAGGGATGATGGTAAGACAATAGATAATCAGGGAGCAATT
AGCACGGTTATTTCTCAATTGAAGCGTAAAGAAAGACCGACTTATTCTGGAG
ATTTACAAGCAAAGTGGTTATTAGAGGAAATTCCATATAGCAAAGCGTTAAG
TGCCCAATATTATGCTGAGGATGATGTCGGTAGTGAAGGTTACATCATCTTTA
ACGATGAATTGGCGATTATAGGTGTGAGAGGTACAGAGCCTTATTTTCAGAG
TAAGAAACCACCTGTTGATAACACTAAATTTAAAATTATAAAAGCGGCTTCA
GGTATGGCGGCTGTGATTGCGGATAAGATAGAAAGTGCTACCGATTCACCAG
GAATGAAAGACCTTATTATAACAGACCTTGATGCTGCGCAAATCGCCCCCGA

AGAGTTTGGTGGAACATATGTTCATCGGGGGTTTTACCAATATACGATGGCTT
TGTTATCTTTGATGGAGAAAGATCTAGGGCTACATAAAATAAAAAAATTCTA
TTGTTGTGGCCATAGCTTAGGGGGCGCTGGTGCATTACTTATAAGTGCATTGA
TTAAAGATAGCTATCATCCTCCAGTATTGCGTTTATATACATATGGAATGCCA
CGAGTGGGAACTCGCAGTTTTGTAGAGCGCTATCAAAATATTTTGCATTACCG
TCATGTTAATAATCACGATTTGGTGCCACAGATACCAACAGTATGGATGAAC
ACAGATGTATCTGAAGGATTCCATGTACTTGATGTATTTAAAAGCCGTGTTGA
TCTGATGAGAAAAATGCTCACTGATGATGATGATGACAATTATCAACACCAT
GGTCATTTGTCTCAGTTACTAACTTATAATTCAAATAACCAAGTCTTACTGAC
ACCCAAACAAACTCAAGTCACTATGCTAGACCTAGCAAATTTAGCCACTAAT
GATTCTGTTGCTATGGTAGATGGTTTATCAGATGCCTCAATAGTAGAGCATGG
AATGGAGCAGTATATACCCAATTTATTTGAGCAGTTAACAGCACTAAGTGAT
GAAAGTTTAATGGTGCATTATCAACGAGCGATTTCAGCCTTAGAGCAGGAAA
TTGCTACGTTGCAACAAAGTTATCTTACAGTTAAACAAGCTTGGATTGAATCT
ATAGGAAATGGCACGCCGACTATGAATATCGGACGTTTGATGAGTGAAATGC
ATTCTATAAATAAATTAATAGAGAACCGTAATAAAATTAGAGGTGAGCTAAG
GCAGATAGTCTCTGATCCTCAGCGTATGCCTGCAACTAAATTTTTAATTTCAC
AGCAAACCTTACCAGATGAAATCAAGGTGCAAATAAGATGA

Appendix A

Portions of this chapter have been published in

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. The Journal of Biological Chemistry 288, 7618-7625.

The animal experiment shown in Figure 9-8 was performed with Verena Bachmann and Ben Kostiuk.

9 Appendix A

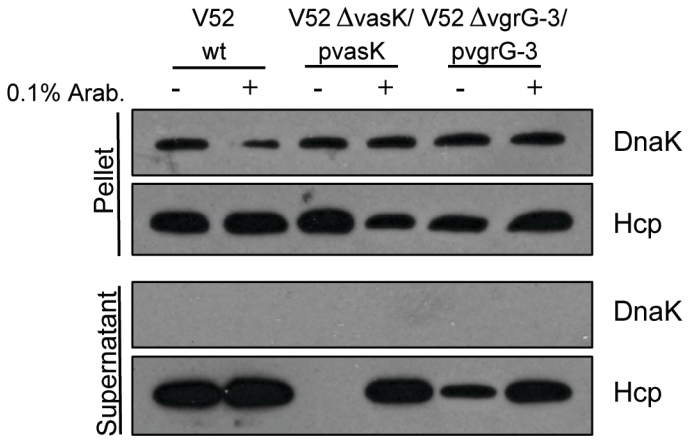


Figure 9-1 Lack of VgrG-3 results in less Hcp secretion of *V. cholerae* strain V52.

Mid-logarithmic cultures were pelleted and the supernatant TCA precipitated for immunoblot analysis. Samples were adjusted for cell density and blotted for DnaK (loading and lysis control) and Hcp. Adapted from (Brooks et al., 2013). This research was originally published in The Journal of Biological Chemistry. © The American Society for Biochemistry and Molecular Biology.

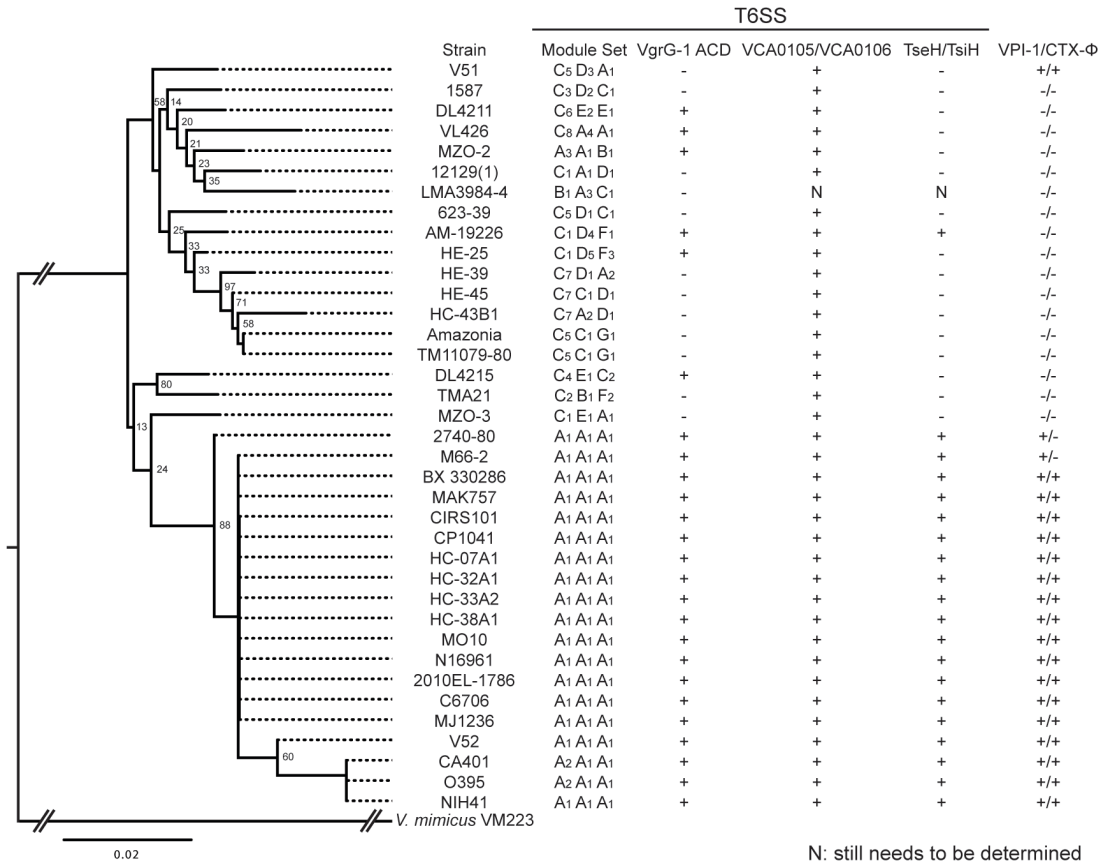
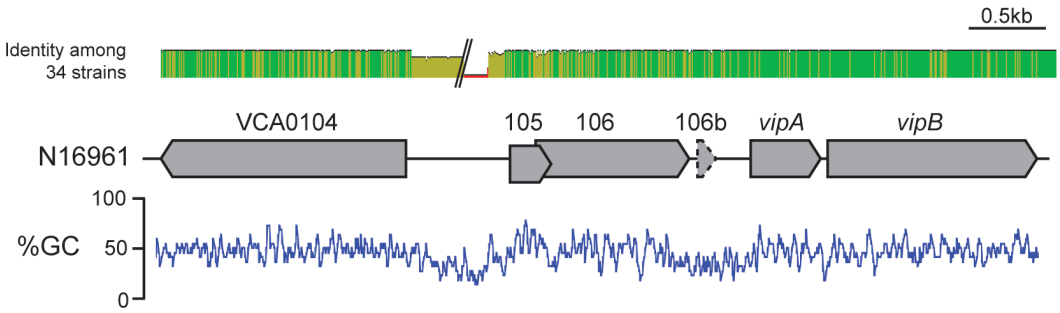


Figure 9-2 Overview over characterized and putative T6SS effector modules

The presence and absence of the ACD, the effector module containing *tseH*/*tsiH* and the presence of the genes VCA0105 and VCA0106 were added to the graph shown in chapter 4. Parts of this figure are adapted from (Unterweger et al., 2014).

A



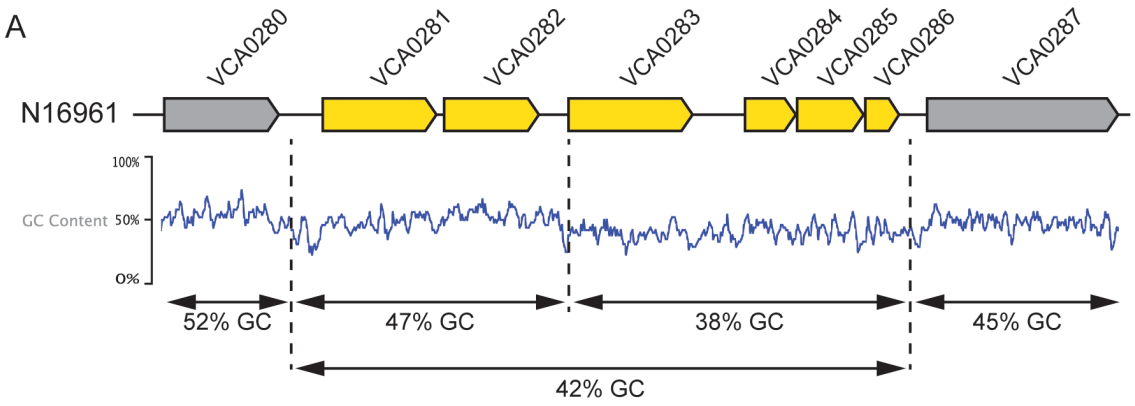
B

Locus Tag	Annotation	Feature
VCA0105	hypothetical protein	PAAR-like superfamily
VCA0106	hypothetical protein	TPR multidomains, SLR repeats
VCA0106b	unannotated	-

Figure 9-3 VCA0105 and VCA0106 are conserved among *V.cholerae* strains

(A) Alignment of the nucleotide sequences of VCA0105 and VCA0106 and surrounding regions of 34 *V. cholerae* strains. The identity among the sequences is shown as bar graph, green bars indicate 100% identity, yellow bars 30-99% identity and red bars less than 30% identity at the indicated positions. The GC content is indicated below a graphical depiction of the open reading frames located at this site. (B) Feature of the proteins encoded in the indicated genes based on a BLAST search.

A



B

Locus Tag	Annotation	Feature
VCA0281	Putative integrase	-
VCA0282	IS5 transposase	DUF772, COG3039, DDE_Tnp_1_6, DDE_Tnp_1
VCA0283	Hypothetical protein	-
VCA0284	Hypothetical protein	PAAR superfamily
VCA0285	TseH	Homologous to proteins with YD-repeat and Rhs repeat-associated core domain
VCA0286	TsiH	Immunity protein

C

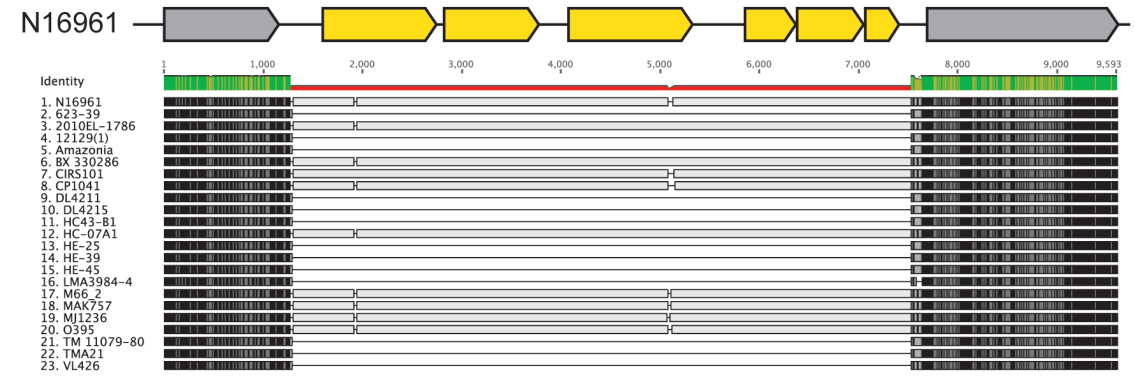
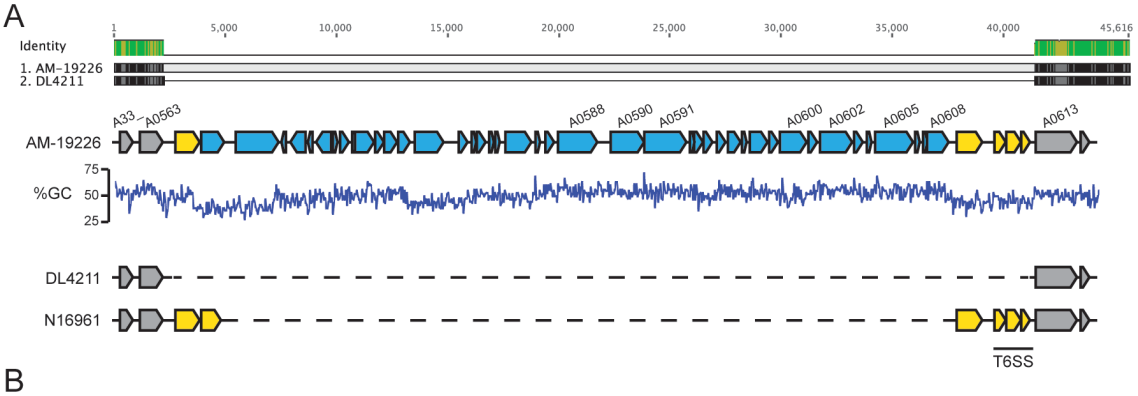


Figure 9-4 VCA0105 and VCA0106 are present or absent among *V. cholerae* strains

(A) Graphical depiction of *tseH* (VCA0285), *tsiH* (VCA0286) and the surrounding genes. The average GC content over the indicated regions is shown. (B) Annotation and features of six genes with the indicated locus tags. Features were retrieved from BLAST. (C) Graphical depiction of the genes shown in (A). The nucleotide sequences over this region were aligned using the MAUVE algorithm. The presence of VCA0281-VCA0286 is indicated by grey boxes, the absence by a grey line.

A



B

Locus Tag	Annotation	Feature
A33_A0588	Phage terminase large subunit GpA	Terminase_GpA superfamily
A33_A0590	phage portal protein lambda family	Phage_portal superfamily
A33_A0591	peptidase U35, phage prohead HK97: Phage major capsid protein HK97	Peptidase U35 superfamily, Phage capsid superfamily
A33_A0600	Phage tail fiber protein	-
A33_A0602	phage tail sheath protein	Phage_sheath_1 superfamily
A33_A0605	conserved hypothetical protein	Cog5412
A33_A0608	Phage protein D	Phage_GPD superfamily

Figure 9-5 AM-19226 contains a unique insertion of 43 genes upstream of *tseH*

(A) Alignment of the nucleotide sequences of AM-19226 and DL4211 using the MAUVE algorithm. Graphical depticion of the open reading frames encoded over the indicated region in the genome of the strain AM-19226. For comparison, the open reading frames of AM-19226 and N16961 are indicated below. The genes encoding the PAAR domain protein, TseH and TsiH are indicated with "T6SS" below the ORFs of N16961. Genes unique to AM-19226 are coloured in blue. Genes that are found in AM-19226 and N16961 but not DL4211 are coloured in yellow. (B) Insertion of genes in AM-19226 is likely derived from a phage. Annotation and feature of the protein encoded in the indicated genes were retrieved from BLAST.

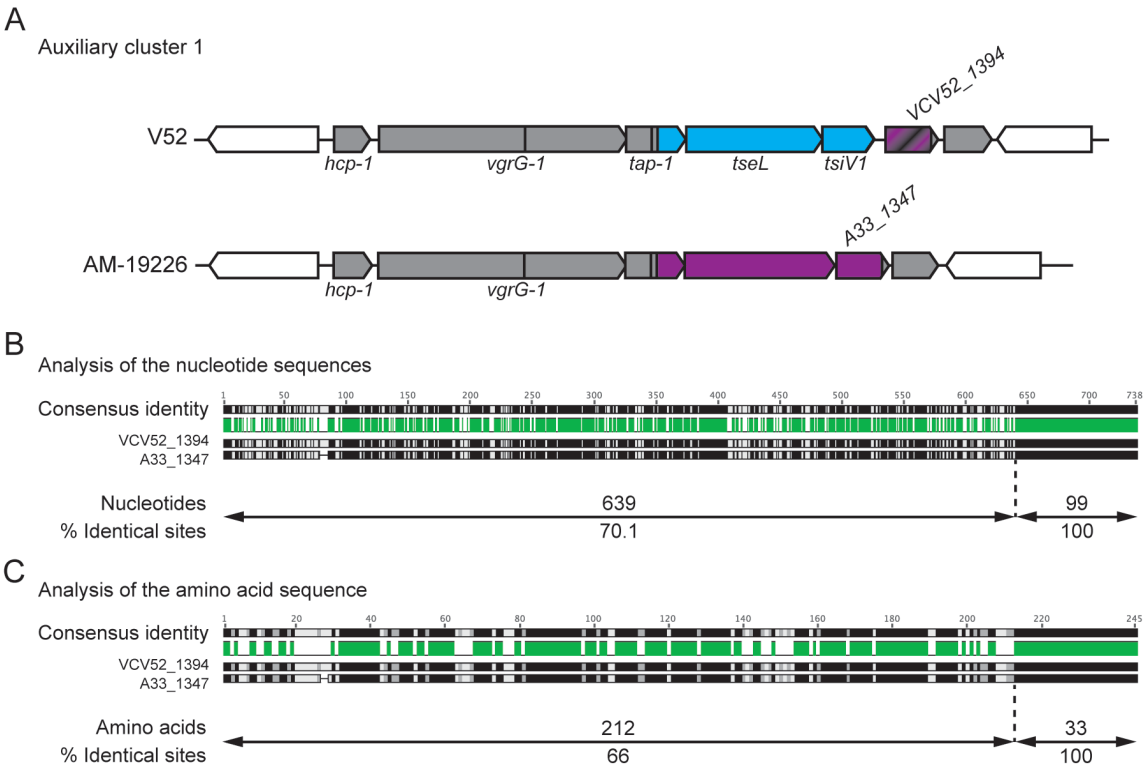


Figure 9-6 Indications for an orphan immunity protein encoded in a gene in auxiliary cluster 1

(A) Cartoon of auxiliary cluster 1 of the strains V52 and AM-19226. The gene name is indicated below, the locus tag on top of the depicted open-reading frames. The locus tag VCV52_1394 corresponds to VC1420 in the strain N16961. The effector modules of the three genes encoding adaptor, effector, immunity protein are highlighted in the same colour. VCV52_1394 and A33_1347 are depicted in similar but not identical colour because, as shown in B, they are homologues but not identical. (B) MUSCLE alignment of the indicated nucleotide sequences. Percent identity over the indicated length is shown below. (C) MUSCLE alignment of the indicated amino acid sequences. Percent identity over the indicated length is shown below.

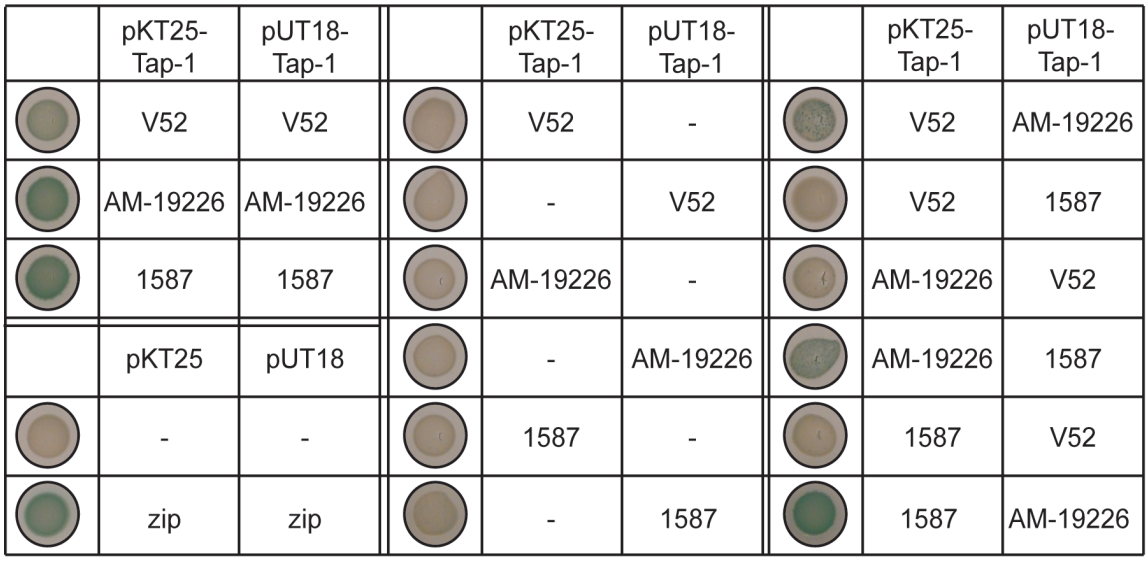


Figure 9-7 Homo- and heterodimers of Tap-1.

Tap-1 from V52, AM-19226 and 1587 form homodimers (top left). If two Tap-1 proteins from different strains share at least one of the two domains in common, they form heterodimers (right column). Results of the bacterial adenylate cyclase two-hybrid (BACTH) assay are shown. *E. coli* strain BTH101 was transformed with the plasmids indicated in the top row. Plasmid pKT expresses Tap-1 from the indicated strains with a N-terminal translational fusion to the T25 fragment of the adenylate cyclase. Plasmid pUT expresses Tap-1 from the indicated strains with a C-terminal translational fusion to the T18 fragment of the adenylate cyclase. Increasing levels of cAMP drive the expression of *lacZ*. Overnight cultures were spotted onto LB plates supplemented with X-Gal (40µg/ml), kanamycin (50µg/ml), ampicillin (100µg/ml), 0.5mM IPTG, and incubated at 30ºC. All spots were incubated on the same plate and the circles were cut out of one picture of the whole plate for clarity. One representative experiment of three experiments, each performed in triplicates, is shown.

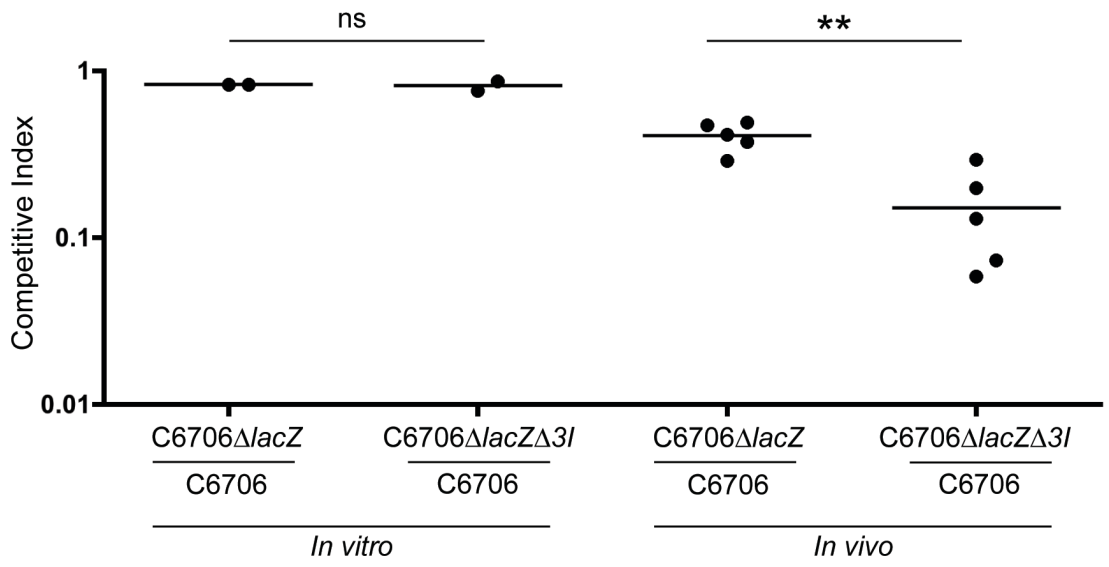


Figure 9-8 Wild-type C6706 outcompetes C6706Δ*tsiV1*Δ*tsiV2*Δ*tsiV3* (C6706Δ3I) under conditions that activate the T6SS

Competition experiments of the indicated strains in liquid culture *in vitro* or in the infant mouse *in vivo* suggest a competitive advantage of wild-type C6706 over C6706Δ3I under conditions that activate the T6SS. The competitive index was determined by dividing the ratio of the two strains after co-incubation, by the ratio of the two strains prior to co-incubation. (A) For the *in vitro* analysis, *V. cholerae* C6706 was mixed with the indicated mutant at a ratio 1:1 and incubated over night in LB broth on the roller drum at 37°C (Taylor et al., 1987). (B) For the *in vivo* analysis, the same strains were mixed and administered orally to infant mice (Ma and Mekalanos, 2010). Mice were kept over night at 30°C and sacrificed the following day. The distal 1cm piece of the small intestine was homogenized, serially diluted and plated on LB plates containing Streptomycin and X-gal. Blue and white colonies were enumerated. A two-tailed student's t-test was performed to determine statistical significance (** $p<0.005$). Each dot represents one mouse.

Appendix B

The herein described screen has been performed together with Nicole Acosta from the laboratory of Tracy Raivio (Department of Biological Sciences, University of Alberta).

10 Appendix B: Keio screen for *E. coli* mutants resistant to T6SS-mediated killing

10.1 Introduction

V. cholerae uses its T6SS to kill bacteria of the same and other species (MacIntyre et al., 2010; Unterweger et al., 2012). *E. coli* is among the bacterial species that are killed by *V. cholerae* in a T6SS-dependent manner. When a mixture of 10^8 O37 serogroup strain *V. cholerae* V52 and 10^7 *E. coli* MG1655 was incubated for four hours at 37 ºC on a LB agar plate, only 1,000 *E. coli* survivors were recovered (MacIntyre et al., 2010). No killing was observed and ~ 10^8 survivors were recovered in a mixture of 10^8 V52Δ*vasK* with a dysfunctional T6SS and 10^7 *E. coli* MG1655 incubated for four hours at 37 ºC on a LB agar plate (MacIntyre et al., 2010). These experiments indicated that *V. cholerae* V52 reduces the number of *E. coli* MG1655 during a four-hour long incubation on an LB agar plate by 100,000 fold (MacIntyre et al., 2010) and that *V. cholerae* V52 lacking *vasK* does not kill *E. coli*. When we conducted this screen in 2010, the molecular mechanism by which *V. cholerae* attacks *E. coli* and the factors *E. coli* succumbs to were unknown. Stefan Pukatzki, Tracy Raivio, Nicole Acosta, and I set out to advance our understanding of T6SS-mediated killing of *E. coli* by identifying *E. coli* mutants that are resistant to a T6SS-mediated attack. The Keio collection consists of 3985 nonredundant, in-frame deletion mutants of nonessential genes in *E. coli* K-12 BW25113 (Baba et al., 2006). This collection provides a powerful tool to look for *E. coli* mutants resistant to a T6SS-mediated attack.

10.2 Results

T6SS-mediated killing of *E. coli* is analyzed in the lab in killing assays in which *V. cholerae* and *E. coli* are mixed at a 10:1 ratio, spotted onto an LB agar plate, incubated at 37 ºC for four hours, scraped off the agar plate, serially diluted, and plated onto LB plates supplemented with antibiotics to separately determine the CFU of *V. cholerae* and *E. coli* after coincubation on LB agar. This experimental setup was not feasible to use for a screen of 3985 *E. coli* mutants for resistance against T6SS-mediated killing. Therefore, we developed a modified protocol of the killing assay based on 96 well plates that is described in detail in the following paragraph.

In the high-throughput screen, *V. cholerae* V52 and the *E. coli* mutants were grown overnight in liquid in a 96 well plate on a shaker at 37 ºC. The next day, the cultures were diluted 1/100 and grown to an OD (600 nm) of 0.5. Each *E. coli* mutant was then mixed with V52 at a ratio 1:1. 2 µl of these mixed cultures were spotted into a 96 well plate filled with 150 µl of LB agar per each well. These 96 well plates were incubated for four hours at 37 ºC standing. After the incubation time, the bacteria were washed off the agar and transferred to a new 96 well plate filled with LB supplemented with kanamycin, to which *E. coli* is resistant and *V. cholerae* is sensitive. The OD (600 nm) of the cultures in the 96 well plate was determined before the plate was incubated for six hours at 37 ºC shaking. After this incubation period, the ODs at 600 nm of the cultures in the 96 well-plate were determined again. The OD_{600} value determined before the 6 hour long incubation time (t0) was subtracted from the OD_{600}-value determined after the 6 hour long incubation time (t6). *E. coli* mutants resistant to a T6SS-mediated attack were expected to grow during the six hours whereas *E. coli* mutants sensitive to a T6SS-mediated attack were expected to show only minimal growth during the six hours. After completing the screen, the data were analyzed to identify resistant Keio mutants. No mutants could be detected that were resistant to a T6SS-mediated attack of V52 (Table 2). From today's perspective, this result can be explained in multiple ways. One reason for the inability to detect a single mutant resistant to the T6SS could be the T6SS-mediated secretion of multiple effectors TseL, VasX, VgrG-3, and TseH with different antiprokaryotic activities such as lipase activity, pore-formation, and peptidoglycan-degradation (Brooks et al., 2013; Dong et al., 2013; Miyata et al., 2013). Even if a single mutant was resistant to TseL, for example, the other three effectors VasX, VgrG-3, and TseH would still be deadly for the *E. coli* mutant. Another reason for the inability to identify a resistant mutant might be the independence of T6SS effector activity on target cell proteins. VasX is a very powerful antiprokaryotic effector (Miyata et al., 2013) and no protein of the target cell has yet been identified that VasX depends on to insert itself into the inner membrane and disrupt the membrane potential. Therefore, a single gene deletion mutant deficient in any nonessential gene in *E. coli* will likely still get killed by the effector VasX.

Table 2 Results of the Keio screen.

Gene/ locus tag	OD(600) t6-t0	Gene/ locus tag	OD(600) t6-t0	Gene/ locus tag	OD(600) t6-t0	Gene/ locus tag	OD(600) t6-t0	Gene/ locus tag	OD(600) t6-t0
thrA	0.08	rimJ	0.35	rfbA	0.08	ygjM	0.15	yjjB	-0.12
thrB	0.04	yceH	0.01	rfbD	0.05	ygjN	-0.03	yjjQ	0.02
thrC	-0.13	mviM	0.00	rfbB	-0.13	ygjQ	-0.08	bglJ	0.00
yaaX	0.07	flgN	0.16	galF	-0.09	ygjR	-0.89	fhuF	-0.19
yaaA	-0.01	flgM	-0.07	wcaM	0.38	ygjU	0.01	rsmC	-0.12
yaaJ	-0.03	flgM	-0.09	wcaL	0.03	ygjV	-0.04	holD	0.01
talB	-0.15	flgA	0.07	wcaK	-0.06	uxaA	-0.01	rimI	-0.11
mog	-0.15	flgB	0.05	wzxC	0.05	uxaC	-0.01	yjjG	0.02
yaaH	0.05	flgC	-0.13	wcaJ	0.00	exuT	0.26	osmY	0.03
yaaW	-0.03	flgD	-0.11	manB	-0.10	exuR	-0.02	yjjU	0.07
yaaI	0.02	flgE	-0.09	manC	0.00	yqjA	-0.05	yjjV	0.02
dnaK	0.01	flgF	-0.20	wcaI	0.13	yqjB	-0.21	yjjW	0.03
dnaJ	-0.01	flgG	-0.04	nudD	0.03	yqjD	-0.02	yjjI	0.04
nhaA	0.01	flgH	0.05	fcl	-0.05	yqjE	0.03	deoA	0.03
nhaR	0.22	flgI	-0.03	gmd	-0.05	yqjK	0.07	deoB	0.14
rpsT	-0.03	flgJ	-0.01	wcaF	-0.05	yqjF	-0.01	deoD	0.10
ileS	-0.02	flgK	-0.01	wcaE	-0.04	yqjG	-0.11	yjjJ	-0.27
fkpB	0.10	flgL	0.07	wcaD	0.06	yhaH	-0.05	lplA	0.08
rihC	0.32	rluC	-0.32	wcaC	0.01	yhaI	-0.05	smp	0.26
carA	0.00	yceF	0.04	wcaB	0.06	yhaJ	0.08	serB	0.08
carB	0.02	yceD	0.00	wcaA	0.02	yhaK	0.04	radA	0.01
yaaV	0.04	rpmF	-0.08	wzc	0.04	tdcD	-0.07	nadR	0.00
caiF	0.43	fabH	-0.01	wzb	-0.01	tdcC	-0.05	yjjK	0.06
caiD	0.49	fabF	-0.03	wza	0.00	tdcB	-0.01	slt	0.02
caiC	0.01	pabC	0.01	yegH	-0.06	tdcA	0.25	trpR	-0.07
caiB	-0.02	yceG	0.01	yegH	0.05	yhaB	0.01	gpmB	0.17
caiA	0.00	ycfH	-0.03	asmA	0.05	yhaC	-0.05	rob	-0.13
caiT	0.67	ptsG	0.06	dcd	0.03	garK	0.09	creA	0.02
fixA	0.01	fhuE	-0.02	udk	-0.01	garL	0.07	creB	0.06
fixB	0.56	ycfF	-0.06	yegE	-0.04	garP	0.00	creC	0.04
fixC	0.44	ycfL	0.06	yegE	0.02	garD	0.04	creD	-0.14
fixX	0.10	ycfM	0.00	alkA	-0.05	sohA	-0.01	arcA	0.04
yaaU	0.16	ycfN	-0.75	yegD	-0.11	yhaV	0.21	yjjY	-0.02
yabF	0.13	nagZ	-0.18	yegI	-0.04	agaZ	0.07	lasT	0.04
kefC	0.15	ycfP	-0.04	yegJ	-0.05	agaV	-0.04	thrL	0.01
apaH	0.21	ndh	0.07	yegK	-0.06	agaW	0.19	htgA	-0.04
apaG	-0.06	ycfJ	-0.03	yegL	-0.01	agaS	-0.07	htgA	-0.30
ksgA	0.03	ycfQ	0.04	yegM	-0.02	agaY	0.00	hokC	0.04
pdxA	0.05	ycfQ	-0.02	yegN	-0.33	agaB	0.10	yaaY	-0.02
surA	-0.25	ycfR	0.00	yegO	-0.02	agaC	0.01	caiE	-0.01
djlA	-0.10	JW1099	-0.04	yegO	-0.03	agaD	0.43	yabI	0.01
yabP	0.18	JW1099	0.02	yegB	0.01	agaI	0.13	leuB	-0.09
rluA	-0.09	mfd	-0.09	yegB	0.04	yraH	0.01	leuB	-0.06
hepA	-0.03	ycfT	0.07	baeS	0.11	yraI	0.10	yacG	0.08
polB	0.02	ycfX	0.05	baeR	0.00	yraJ	-0.05	hpt	0.05
araC	0.18	ycfX	0.06	yegP	0.03	yraK	-0.01	yadD	0.09
yabI	0.05	cobB	0.08	ogrK	-0.08	yraM	-0.02	ligT	-0.04
thiQ	-0.05	cobB	0.09	JW2068	0.00	yraN	0.09	ligT	-0.31
thiP	-0.02	ymfA	0.00	yegR	-0.07	yraO	0.33	eriC	0.03
tbpA	-0.02	potD	0.03	yegS	0.05	yraP	0.20	cdaR	0.03
yabN	0.17	potC	0.03	yegS	0.00	yraQ	0.04	yaeI	-0.03
setA	0.01	potB	-0.17	gatR	0.26	yhbO	-0.06	JW5015	-0.01
leuD	0.05	potA	0.59	gatD	-0.18	yhbP	0.25	yaeF	0.01
leuC	0.12	pepT	-0.05	gatC	0.02	yhbQ	0.36	yafD	0.02
leuA	0.03	ycfD	0.04	gatB	0.00	yhbS	0.39	mltD	-0.04
leuL	-0.16	phoQ	-0.09	gatZ	-0.04	yhbT	-0.06	yafV	-0.04
leuO	0.22	phoP	0.05	fbaB	0.02	yhbU	0.01	fadE	0.11

ilvI	0.08	trmU	-0.16	fbaB	-0.04	yhbW	-0.07	mbhA	0.03
ilvH	0.06	ymfB	0.00	yegT	0.00	mtr	0.09	yafX	-0.02
fruR	0.19	ymfB	-0.05	yegU	0.07	deaD	0.01	ykfF	0.05
mraZ	0.08	ymfC	0.08	yegV	0.05	nlpI	-0.01	JW5025	-0.08
mraZ	0.13	icd	-0.08	yegW	-0.09	JW3133	0.03	mmuP	-0.02
mraW	0.04	ymfD	0.14	yegX	0.01	rpsO	0.00	mmuP	-0.10
mutT	0.00	ymfE	0.01	thiD	0.00	truB	0.05	JW5028	0.03
yacG	0.04	lit	0.08	thiM	0.02	rbfA	0.03	JW5029	-0.01
yacF	-0.05	intE	0.33	yohL	0.08	argG	-0.19	JW5029	-0.01
coaE	0.07	ymfG	0.49	yohM	-0.09	secG	0.53	yagV	0.09
guaC	0.17	ymfH	0.61	yohN	-0.05	mrsA	0.06	ykgK	0.05
hofC	0.48	ymfI	-0.15	yehA	0.04	folP	-0.06	JW5032	-0.03
hofB	0.02	ymfJ	0.01	yehB	0.02	rrmJ	0.06	JW5032	-0.05
ppdD	-0.04	ymfL	0.01	yehC	-0.02	yhbY	-0.06	ykgL	0.06
nadC	0.01	ymfM	0.01	yehD	-0.14	greA	-0.03	JW5034	-0.01
ampD	0.12	ymfN	0.00	yehE	0.10	dacB	0.02	JW5034	-0.03
ampE	-0.08	ymfR	0.01	mrp	0.04	yhbE	0.19	rpmE2	-0.07
aroP	-0.16	ymfO	-0.03	yehI	-0.02	sfsB	-0.25	rpmE2	-0.12
pdhR	0.29	ymfP	0.10	yehK	-0.03	yrbA	-0.35	ykgA	0.04
aceE	0.05	ymfQ	-0.01	yehL	-0.02	yrbC	0.02	ykgB	-0.02
aceF	-0.16	ycfK	0.01	yehM	0.01	yrbD	-0.12	ykgI	0.09
lpdA	-0.01	ycfK	-0.05	yehP	-0.03	yrbE	0.00	ykgC	-0.06
yacH	0.04	tfaE	0.14	yehQ	0.02	yrbF	0.10	ykgC	-0.03
acnB	0.13	stfE	0.00	yehR	0.01	yrbG	0.06	ykgE	0.03
yacL	0.09	pinE	-0.09	yehS	0.12	yrbH	0.01	ykgG	0.05
speD	0.02	mcrA	0.05	yehT	0.04	yrbI	0.00	yahH	0.08
speE	0.00	JW5173	0.01	yehT	0.07	yhbG	0.14	yahM	0.05
yacC	0.00	elbA	0.05	yehU	0.06	rpoN	0.05	cynR	0.07
cueO	-0.02	ycgX	-0.03	mlrA	0.04	yhbH	0.15	mhpT	0.21
gcd	0.10	ycgE	-0.07	yehW	-0.02	ptsN	-0.11	JW5047	-0.10
yadG	-0.03	ycgF	-0.02	yehX	-0.02	yhbJ	-0.02	JW5049	-0.03
yadH	-0.03	ycgF	0.01	yehY	-0.55	ptsO	-0.05	JW5049	-0.01
yadI	0.00	ycgZ	0.07	yehZ	-0.02	yrbL	0.02	yaiU	0.07
yadE	0.02	ymgA	0.23	bglX	-0.19	mtgA	0.02	ampH	-0.01
panD	-0.08	ymgB	0.06	dld	0.02	elbB	-0.02	yaiZ	0.06
yadD	-0.01	ymgC	-0.08	yohC	0.00	arcB	-0.02	psiF	-0.05
panC	0.03	ycgG	-0.03	yohD	-0.09	yhcC	0.15	psiF	-0.03
panB	-0.11	ymgF	-0.03	yohF	-0.18	gltB	0.21	proY	-0.05
yadC	-0.03	ydeU	0.06	yohG	-0.03	gltD	0.14	yajI	0.02
yadK	0.05	ymgD	-0.02	yohG	-0.02	gltF	-0.05	thiJ	-0.01
yadL	0.04	JW5178	-0.02	yohI	-0.02	yhcA	0.10	yajQ	0.07
yadM	-0.03	JW5179	-0.01	yohJ	0.05	yhcD	0.23	yajR	0.09
htrE	0.11	JW1162	0.06	yohK	0.08	yhcE	0.02	bolA	0.04
ecpD	0.06	minC	0.04	cdd	0.08	yhcF	0.21	mdlB	-0.07
yadN	-0.07	ycgJ	0.02	sanA	0.04	yhcG	0.30	mdlB	-0.02
pcnB	0.03	ycgK	-0.01	yeiT	-0.05	yhcH	0.08	ylaB	0.06
yadB	-0.33	ycgL	0.00	yeiA	-0.05	nanE	0.01	ylaC	0.07
dksA	-0.49	ycgM	-0.03	mglC	-0.15	nanT	0.06	JW5064	0.02
sfsA	-0.14	ycgM	0.10	mglA	0.43	nanA	0.03	ybbM	0.09
hrpB	-0.02	ycgN	-0.02	mglB	-0.09	nanR	0.19	ybbN	0.05
mrcB	-0.04	hlyE	-0.03	galS	0.27	dcuD	0.04	JW5068	-0.05
fhuA	0.03	umuD	-0.02	yeiB	-0.10	sspB	0.05	JW5069	0.01
fhuC	-0.07	umuC	0.03	yeiG	0.04	sspA	-0.02	ybcJ	0.04
fhuD	0.00	nhaB	0.00	cirA	-0.09	yhcM	0.03	sfmH	0.00
fhuB	0.14	fadR	0.48	lysP	-0.11	degQ	-0.01	sfmF	-0.08
eriC	-0.11	ycgB	-0.02	yeiE	-0.01	mdh	-0.13	sfmF	-0.04
yadS	0.12	dadA	-0.08	yeiH	0.07	argR	-0.06	fimZ	0.10
btuF	0.05	dadX	-0.07	nfo	-0.04	yhcO	-0.01	JW5075	-0.05
mtn	-0.11	ldcA	-0.04	yeiI	-0.02	yhcP	-0.04	JW5075	-0.15
dgt	0.36	mltE	-0.09	yeiI	-0.01	yhcQ	0.08	ylcG	0.08
degP	-0.02	ycgR	-0.03	yeiJ	0.07	yhcS	0.02	nmpC	-0.02
cdaR	0.00	ymgE	-0.12	rihB	0.12	tldD	0.05	rzpD	0.02
yaeH	-0.02	ycgY	0.02	nsr	-0.01	yhdP	-0.02	rzoD	0.02
yaeI	0.00	treA	0.01	yeiM	0.06	cafA	0.04	ybcV	0.07
yaeI	-0.01	ycgS	-0.03	yeiC	-0.25	yhdE	0.00	cusS	0.09
glnD	-0.07	ycgT	-0.02	fruA	-0.13	yhdA	0.03	ybdF	0.06

hlpA	-0.04	ycgV	0.06	fruK	0.17	yhdH	-0.01	hokE	0.02
rnhB	0.28	ychF	-0.01	fruB	-0.24	yhdT	-0.01	fepA	0.04
ldcC	0.19	ychH	-0.02	setB	-0.13	panF	-0.03	citF	-0.06
yaeR	-0.02	ychM	0.02	yeiW	0.29	prmA	0.01	lipB	0.07
rof	0.00	sirB2	-0.01	yeiQ	-0.04	yhdG	-0.03	ybeB	0.03
yaeP	0.02	sirB1	0.01	yeiR	-0.04	fis	-0.09	ybeQ	-0.03
yaeQ	-0.02	chaA	0.01	yeiU	-0.08	yhdU	-0.12	gltI	-0.04
yaeJ	-0.09	chaB	0.07	spr	0.04	envR	0.65	gltI	-0.10
cutF	-0.03	chaC	0.01	rtn	-0.05	acrE	0.09	ybfG	-0.03
yaeF	-0.11	ychN	0.05	yejA	-0.05	acrF	-0.03	ybfG	-0.06
yaeB	0.07	ychN	0.05	yejB	-0.12	yhdV	0.03	kdpE	0.09
rcsF	0.14	ychO	0.07	yejE	-0.36	yhdW	-0.05	abrB	0.03
metQ	0.13	narL	0.05	yejF	-0.03	yhdZ	0.04	ybgO	0.05
metI	0.26	narX	0.21	rsuA	0.02	yrdB	0.02	ybgQ	0.05
metN	-0.01	narK	0.01	yejH	0.02	aroE	-0.08	tolB	-0.01
yaeD	-0.03	narG	0.02	rplY	0.32	smg	1.12	tolB	-0.05
dkgB	0.01	narH	0.08	yejK	-0.12	sun	0.20	JW5101	0.00
yafC	0.28	narJ	0.70	yejL	0.09	trkA	-0.09	ybhT	0.09
yafD	0.01	narI	0.11	JW2177	0.00	mscL	0.00	ybhJ	0.04
yafE	-0.04	tpr	-0.14	narP	0.07	yhdL	-0.12	ybhF	-0.01
gloB	0.12	purU	-0.34	ccmH	-0.03	zntR	0.22	ybhF	-0.15
yafS	0.04	ychJ	0.08	dsbE	-0.01	yhdN	0.04	ybiX	0.08
rnhA	0.11	ychJ	0.11	ccmF	-0.03	rpmJ	-0.37	ybiM	-0.01
dnaQ	0.05	ychK	0.01	ccmE	0.00	pioO	-0.08	ybiN	-0.03
yafT	0.06	hnr	0.02	ccmD	0.04	gspA	0.06	ybiO	0.00
yafU	-0.01	galU	0.68	ccmC	0.09	gspC	0.34	fsaA	0.11
yafV	0.02	hns	0.26	ccmB	0.09	gspD	-0.07	yliA	0.06
ivy	-0.01	tdk	-0.18	napC	-0.08	gspE	0.61	yliB	0.02
fadE	0.06	JW1227.5	-0.01	napH	0.21	hofF	-0.13	ybjG	-0.02
lpcA	-0.06	adhE	-0.09	napG	-0.18	hofG	0.11	ybjI	0.04
yafJ	0.00	ychE	-0.07	napA	0.20	hofH	0.72	ybjK	0.05
yafK	0.04	oppA	0.05	napD	0.04	gspJ	0.22	JW5115	-0.02
yafQ	0.09	oppB	0.04	napF	0.44	gspK	0.19	ybjT	0.02
dinJ	-0.04	oppC	0.01	eco	0.08	hofD	0.01	hcr	-0.04
yafL	0.02	oppD	0.02	mqo	0.14	bfr	-0.06	dmsA	-0.07
yafM	0.00	oppF	0.07	yojI	-0.17	bfd	0.02	ycaM	0.02
fhiA	0.02	yciU	-0.03	alkB	-0.01	chiA	0.10	ycaI	0.05
dinB	0.03	cls	0.13	ada	-0.06	tufA	0.04	ssuC	-0.04
yafN	0.27	kch	0.01	ompC	-0.02	yheL	-0.03	ycbQ	0.03
yafO	0.02	yciI	0.02	yojN	-0.15	yheM	-0.02	ycbV	-0.02
yafP	0.07	yciA	0.03	rcsB	-0.10	yheN	-0.05	ycbF	0.08
ykfJ	0.04	yciA	0.01	rcsC	-0.08	fkpA	0.88	ycbW	-0.01
prfH	-0.08	ispZ	-0.01	JW2207	-0.01	slyX	0.02	ycbX	0.07
pepD	-0.02	yciC	-0.04	atoS	-0.11	slyD	0.31	ymbA	-0.05
gpt	-0.02	ompW	-0.05	atoC	-0.02	yheV	0.07	yccS	-0.07
yafA	0.06	yciE	-0.02	atoD	-0.07	kefB	-0.13	mgsA	-0.07
crl	-0.14	yciF	-0.04	atoA	-0.04	yheR	0.12	yccU	0.01
phoE	-0.10	yciG	-0.03	atoE	0.06	yheS	0.01	yccX	-0.03
proB	0.02	trpA	-0.09	atoB	0.00	yheT	0.16	etp	0.04
proA	0.01	trpB	0.13	yfaP	-0.01	yheU	0.05	ymcD	0.02
ykfI	0.00	trpC	-0.25	yfaQ	-0.01	prkB	-0.03	cspH	-0.01
yafW	0.05	trpD	-0.01	yfaS	0.00	yhfA	0.02	torS	0.00
ykfH	-0.03	trpE	-0.01	yfaS	-0.02	crp	-0.10	ymdF	-0.04
ykfG	0.00	trpL	-0.02	yfaT	-0.02	argD	0.06	ycdG	0.09
yafX	-0.01	trpH	0.08	yfaA	-0.02	pabA	0.04	ycdH	0.00
ykfF	0.02	yciQ	0.02	ubiG	0.04	fic	0.30	ycdL	0.03
ykfB	-0.06	yciL	0.03	ubiG	-0.06	yhfG	-0.03	JW5140	-0.05
yafY	-0.02	btuR	-0.10	yfaL	0.05	ppiA	0.32	ycdN	-0.29
ypjK	0.01	yciK	-0.04	yfaL	-0.03	tsgA	0.00	ycdR	0.19
yafZ	-0.05	sohB	-0.05	yfaE	0.05	nirB	0.37	ycdT	0.02
ykfA	0.04	yciN	0.00	inaA	0.03	nirD	0.04	JW5145	0.02
perR	0.26	cysB	0.04	glpQ	-0.08	nirC	0.01	ycdZ	-0.02
insM	-0.02	acnA	-0.08	glpT	-0.01	cysG	-0.38	JW5148	-0.07
ykfC	0.06	pgpB	0.03	glpA	0.00	yhfL	0.08	JW5148	0.07
mmuM	0.06	yciS	-0.02	glpB	-0.09	yhfM	-0.13	ymdC	-0.03
afuC	0.11	yciM	-0.03	glpC	-0.14	yhfQ	0.03	yceK	0.01

afuB	-0.19	pyrF	0.11	JW2238.5	-0.01	yhfS	-0.06	yceP	0.00
JW0258	-0.03	yciH	-0.04	yfaU	0.03	yhfT	0.05	flgH	-0.01
yagB	-0.11	osmB	-0.09	yfaV	-0.17	php	0.03	yceF	-0.07
yagA	0.02	yciT	0.03	yfaW	0.01	yhfW	0.14	plsX	-0.06
yagE	-0.04	JW1277	-0.02	yfaW	0.12	yhfX	0.12	plsX	-0.22
yagF	-0.05	yciR	0.01	yfaX	0.03	yhfZ	0.30	ycfM	0.03
yagH	0.06	rnb	-0.24	cinA	-0.01	gph	0.01	ycfP	0.03
yagI	-0.06	yciW	-0.01	yfaZ	0.02	rpe	0.17	ycfQ	0.02
argF	0.04	ycjD	0.09	yfaO	0.01	dam	-0.04	ycfS	0.02
yagJ	-0.07	sapF	0.31	yfaO	-0.02	damX	0.30	JW5163	-0.05
yagK	-0.06	sapD	0.06	ais	0.07	aroB	-0.33	JW5163	-0.17
yagL	0.09	sapC	0.11	yfbE	0.04	hofQ	-0.02	ymfA	0.05
yagM	-0.18	sapB	0.03	yfbE	0.02	yrfB	-0.02	ycfC	-0.02
yagN	0.12	sapA	0.12	yfbF	-0.01	yrfC	0.01	ymfE	0.01
intF	-0.26	ymjA	-0.03	yfbF	-0.02	mrcA	0.08	JW5167	-0.01
yagP	0.06	ycjJ	-0.15	yfbG	-0.06	nudE	0.02	ymfI	0.05
yagQ	0.10	ycjL	0.09	yfbH	0.00	yrfG	0.03	JW5169	-0.04
yagR	-0.03	ycjC	0.01	yfbI	0.02	hslR	0.06	ymfP	0.00
yagR	-0.01	aldH	0.05	yfbI	-0.01	yhgE	0.08	ymfS	-0.04
yagS	0.09	ordL	0.08	yfbW	-0.01	pckA	0.14	stfE	0.07
yagS	0.07	goaG	-0.11	yfbJ	-0.01	envZ	0.03	JW5173	0.08
yagT	-0.03	pspF	0.31	pmrD	-0.02	ompR	0.04	ycgG	-0.03
yagU	0.01	pspA	0.30	menE	0.00	greB	-0.04	ydeU	0.06
yagU	0.00	pspB	0.01	menC	-0.13	yhgF	0.23	ymgD	0.04
ykgJ	-0.01	pspC	0.16	menB	0.01	feoA	-0.13	JW5178	0.04
yagV	-0.01	pspD	0.15	yfbB	-0.01	feoB	0.08	JW5179	-0.03
yagW	0.05	pspE	-0.03	yfbB	-0.13	yhgG	0.01	ycgN	-0.01
yagX	0.17	ycjM	0.05	menF	0.02	yhgA	0.02	hlyE	0.05
yagY	0.03	ycjN	0.00	elaB	0.48	bioH	0.04	dsbB	-0.07
yagZ	0.02	ycjO	0.13	elaA	0.53	yhgI	-0.12	dsbB	0.10
ykgK	0.04	ycjP	-0.06	elaC	-0.26	gntT	-0.02	JW5183	0.10
ykgL	-0.01	ycjQ	0.22	yfbK	0.08	malQ	0.04	ycgO	0.01
eaeH	-0.04	ycjR	0.01	yfbK	-0.10	malP	-0.02	ycgC	-0.03
ykgA	0.04	ycjS	0.04	yfbL	0.01	malT	0.22	ycgS	0.07
ykgB	0.07	ycjT	0.00	yfbL	-0.04	rtcB	0.03	ycgT	0.07
ykgI	-0.22	ycjU	-0.09	yfbN	0.11	rtcR	-0.07	dhaR	-0.10
ykgD	0.09	ycjV	-0.14	yfbO	0.32	glpR	0.19	ychM	-0.03
ykgE	-0.04	ompG	0.05	nuoN	0.03	glpG	-0.01	JW5191	-0.10
ykgF	0.02	ycjW	0.05	nuoM	-0.06	glpE	0.29	JW5194	0.03
ykgG	0.04	ycjX	-0.04	nuoL	-0.05	glpD	-0.24	tonB	-0.08
ykgH	0.09	ycjF	-0.03	nuoK	0.06	yzgL	0.05	yciO	0.00
betA	-0.22	tyrR	-0.07	nuoJ	0.07	glgP	0.38	yciQ	0.04
betB	-0.28	tpx	0.00	nuoI	0.06	glgA	0.09	yciX	0.07
betI	-0.07	ycjG	0.00	nuoH	0.05	glgC	-0.02	yciX	-0.04
betT	0.17	mpaA	-0.05	nuoG	0.03	glgX	-0.02	yciX	-0.31
yahA	0.06	JW1319.5	-0.01	nuoF	0.10	glgB	-0.13	yciW	0.03
yahB	0.35	ycjY	-0.03	nuoE	-0.06	yhgN	0.17	ycjK	-0.04
yahC	0.05	ycjZ	0.04	nuoB	-0.05	gntK	-0.19	ycjK	-0.11
yahD	-0.05	JW1321.5	0.02	lrhA	0.04	yhhW	0.02	ycjR	-0.02
yahE	0.03	mppA	-0.03	yfbQ	-0.01	yhhX	-0.03	JW5203	-0.02
yahF	0.01	ynaI	-0.11	yfbR	-0.02	yhhY	0.06	JW5203	-0.42
yahG	-0.05	ynaJ	0.03	yfbS	0.01	yhhZ	0.08	abgA	0.05
yahH	0.02	ydaA	0.08	yfbS	-0.23	yrhA	-0.11	ydaM	-0.02
yahI	0.15	fnr	-0.05	yfbT	0.08	yrhB	0.00	ydaQ	0.07
yahJ	-0.06	ogt	-0.01	yfbT	-0.02	ggt	0.04	lar	0.01
yahK	-0.16	abgT	0.01	yfbU	-0.06	yhhA	0.04	sieB	-0.10
yahL	0.03	abgB	-0.02	yfbV	0.13	ugpQ	0.01	sieB	-0.17
yahM	0.25	abgA	-0.03	ackA	0.03	ugpC	0.22	ydaG	-0.01
yahN	-0.04	abgR	0.04	pta	0.03	ugpE	0.42	ydaW	0.05
yahO	0.07	ydaL	0.02	yfcC	0.06	ugpA	-0.03	rzpR	-0.02
yahO	-0.03	ydaM	0.02	yfcC	-0.09	ugpB	-0.04	rzoR	-0.02
prpR	0.01	ydaM	0.02	yfcD	-0.02	livF	0.03	ydbJ	0.04
prpB	0.13	ydaN	0.03	yfcE	-0.02	livG	-0.01	ydbL	-0.02
prpC	0.21	ydaN	-0.02	yfcE	0.03	livM	-0.03	paaD	0.03
prpD	-0.04	JW1336.5	0.02	yfcF	0.10	livH	0.09	paaK	-0.07
prpE	0.04	dbpA	-0.04	yfcG	-0.01	livK	-0.04	paaK	-0.21

codB	0.07	ydaO	0.02	yfcG	-0.01	yhhK	0.03	ydbD	-0.16
codA	-0.14	intR	0.23	folX	0.02	livJ	-0.04	hrpA	0.04
cynR	-0.08	ydaQ	-0.14	yfcH	0.02	yhhF	0.05	cybB	0.06
cynT	0.05	ydaC	-0.02	yfcH	-0.12	yhhM	0.07	hokB	-0.03
cynS	0.09	recT	0.00	yfcI	-0.02	yhhN	0.02	hokB	-0.09
cynX	0.06	recE	0.44	hisP	0.00	zntA	0.02	ydcI	0.04
lacA	0.02	racC	-0.04	hisM	0.00	sirA	-0.12	JW5227	0.01
lacY	0.00	ydaE	-0.03	hisQ	-0.01	acpT	-0.12	ydcM	0.09
lacI	-0.63	kil	-0.04	hisJ	0.07	nikA	-0.02	ydcO	0.10
mhpR	0.22	ydaF	-0.05	argT	0.00	nikB	0.03	JW5230	-0.07
mhpA	0.02	ydaG	0.06	ubiX	0.04	nikC	0.24	JW5231	-0.05
mhpB	0.07	ydaS	-0.02	purF	0.13	nikD	0.06	JW5231	0.03
mhpC	-0.37	ydaT	-0.10	cvpA	0.04	nikE	-0.04	ydcX	0.07
mhpD	-0.01	ydaU	-0.12	dedD	-0.03	nikR	-0.09	yncA	0.01
mhpF	-0.24	ydaV	0.07	dedA	-0.06	yhhH	0.13	ansP	-0.02
mhpE	0.21	ydaW	0.01	truA	-0.30	yhhI	0.10	yncH	0.00
yaiL	0.10	rzpR	-0.07	usg	-0.01	yhiI	-0.04	JW5236	-0.03
yaiM	-0.31	trkG	0.09	pdxB	0.09	yhiJ	0.06	JW5236	-0.14
adhC	-0.19	ynaK	-0.02	div	-0.05	yhiL	0.06	JW5237	0.01
yaiN	0.00	ydaY	-0.03	yfcJ	-0.03	yhiN	-0.16	sfcA	0.00
yaiO	0.02	ynaA	0.03	yfcJ	-0.02	pitA	-0.24	bdm	0.06
JW0350	-0.02	JW1362	-0.16	yfcL	-0.05	uspB	0.13	yddS	0.10
yaiP	-0.02	JW1365	0.00	yfcM	0.09	uspA	0.11	yddV	0.05
yaiS	0.01	stfR	0.01	yfcA	0.10	yhiP	0.10	yddA	-0.05
tauA	0.04	tfaR	-0.09	aroC	0.00	prlC	0.01	ydeN	0.09
tauB	0.00	pinR	0.08	yfcB	0.03	yhiR	0.18	JW5244	-0.03
tauC	0.06	ynaE	0.09	yfcN	0.09	gor	-0.51	yneE	0.11
tauD	-0.22	ynaF	-0.04	yfcO	-0.07	arsR	-0.09	yneI	0.08
yaiT	-0.01	ompN	-0.03	yfcP	-0.01	arsB	0.02	marR	0.03
yaiU	-0.01	ydbK	0.04	yfcQ	-0.06	arsC	0.05	marA	0.01
yaiV	0.03	ydbJ	-0.05	yfcR	-0.02	yhiS	-0.19	eamA	-0.01
sbmA	-0.05	hslJ	-0.13	yfcS	-0.02	slp	0.00	JW5251	-0.07
yaiW	0.01	ldhA	0.02	yfcU	-0.02	yhiF	0.02	ydfO	0.05
yaiY	-0.07	ydbH	-0.06	yfcV	0.04	hdeA	0.00	gnsB	0.08
yaiZ	-0.04	ynbE	-0.07	sixA	0.04	hdeD	0.10	ynfN	-0.09
ddlA	0.07	feaR	0.08	yfcX	-0.02	yhiE	0.05	essQ	0.06
yaiB	0.02	feaB	0.16	yfcY	-0.10	yhiU	0.00	ydfU	0.07
phoA	0.06	tynA	-0.05	yfcZ	-0.07	yhiV	-0.10	JW5257	-0.13
yaiC	-0.02	maoC	0.22	fadL	0.00	gadW	0.05	ynfC	-0.03
proC	0.05	paaA	0.01	yfdF	0.06	gadX	0.01	ynfD	-0.04
yaiI	-0.02	paaB	0.01	vacJ	0.01	gadA	-0.03	ynfF	-0.11
aroL	0.05	paaC	0.02	yfdC	-0.03	yhjA	0.01	ynfH	-0.08
yaiA	0.02	paaD	-0.01	intS	-0.17	treF	-0.05	ynfI	-0.05
aroM	-0.04	paaE	0.22	yfdG	-0.02	yhjB	0.06	ynfJ	0.03
yaiE	0.02	paaF	0.02	yfdH	0.03	yhjC	0.02	ynfK	-0.03
ykiA	-0.01	paaG	-0.02	yfdH	0.01	yhjD	0.00	ydgJ	0.07
rdgC	0.03	paaH	-0.01	yfdI	0.08	yhjE	-0.03	JW5266	-0.05
yajF	0.02	paaI	0.05	yfdL	0.03	yhjG	0.11	slyA	0.15
yajF	0.27	paaJ	-0.01	xylU	-0.01	yhjH	-0.01	ydhL	0.19
araJ	0.02	paaX	0.05	yfdO	-0.03	dctA	-0.02	JW5269	0.17
sbcC	0.05	paaY	0.01	yfdP	-0.05	bcsZ	0.10	ydhO	0.15
sbcD	-0.16	ydbA	-0.03	yfdQ	-0.01	yhjR	0.02	ydhX	-0.05
phoB	0.07	ydbC	-0.06	yfdR	-0.04	yhjU	0.06	ydhV	0.14
phoR	0.43	ydbD	0.04	yfdS	-0.07	ldrD	0.95	sufB	0.17
brnQ	0.36	ynbA	-0.03	JW5386	-0.02	yhjV	0.59	ydiO	0.01
malZ	-0.51	ynbA	0.17	dsdC	0.23	dppF	0.03	ydiQ	-0.03
yajB	-0.03	ynbB	0.06	dsdX	0.01	dppD	0.04	JW5279	-0.05
queA	-0.01	ynbC	-0.02	dsdA	0.23	dppC	0.07	pfkB	0.19
tgt	-0.06	ynbD	-0.02	emrY	0.00	dppB	0.36	astD	-0.06
yajC	-0.21	acpD	0.03	emrK	0.00	dppA	-0.05	ydjY	0.16
yajD	0.01	hrpA	0.00	evgA	-0.04	yhjX	-0.06	ynjB	0.25
tsx	-0.04	ydcF	0.01	evgS	-0.06	tag	0.11	ynjC	0.18
yajI	-0.07	aldA	0.17	yfdE	-0.01	yiaC	0.78	ynjD	0.27
ybaD	0.08	gapC	-0.05	yfdV	0.00	yiaD	-0.03	ynjE	0.13
ybaD	0.01	cybB	0.07	yfdU	-0.08	yiaG	-0.14	ynjI	0.11
nusB	0.76	ydcA	-0.01	yfdW	0.08	cspA	-0.06	ydjH	0.02

pgpA	-0.08	JW1416.5	-0.02	yfdX	-0.02	hokA	0.04	ydjK	0.10
yajO	-0.17	trg	0.04	ypdI	0.02	glyS	0.08	yeaJ	0.02
xseB	0.35	ydcI	0.00	yfdY	-0.02	JW3532	0.03	yeaP	0.10
thiI	-0.14	ydcJ	-0.10	ddg	-0.03	yiaH	-0.05	yeaV	0.19
panE	0.01	ydcG	0.05	yfdZ	-0.02	yiaA	0.06	yeaW	0.62
yajQ	0.01	JW1421	-0.02	ypdA	0.02	xylB	-0.11	yoaB	0.16
yajR	0.15	ydcH	-0.06	ypdA	-0.02	xylA	0.19	yoaC	0.20
cyoE	0.00	rimL	-0.06	ypdB	-0.06	xylF	0.08	yebN	0.19
cyoD	0.19	ydcK	-0.07	ypdB	-0.22	xylG	-0.24	JW5298	0.03
cyoC	0.20	tehA	0.00	ypdC	0.10	xylH	-0.19	yebQ	0.09
cyoB	0.20	tehB	-0.02	ypdC	-0.05	xylR	-0.47	proQ	0.13
cyoA	0.13	ydcL	-0.10	ypdD	-0.07	bax	0.00	yebU	0.12
ampG	-0.36	JW1427.5	0.00	ypdE	-0.07	malS	0.42	yebV	0.02
yajG	-0.06	ydcM	0.12	ypdG	0.06	avtA	0.04	yebW	-0.03
bolA	0.45	ydcO	-0.10	ypdH	-0.25	ysaA	0.50	yebA	-0.01
tig	-0.09	ydcN	-0.02	yfeO	0.08	yiaJ	0.07	JW5305	-0.16
clpP	-0.02	ydcN	0.04	yfeO	-0.07	yiaK	0.03	yebB	0.12
clpX	0.00	ydcP	-0.06	ypeC	-0.04	yiaL	0.04	yecD	0.16
lon	0.04	yncJ	-0.03	mntH	-0.22	yiaM	0.06	yecN	0.19
hupB	-0.08	ydcQ	-0.10	nupC	-0.03	yiaO	0.02	yecM	0.18
ppiD	0.07	ydcR	0.05	yfeA	-0.04	lyx	0.49	yecT	0.16
ybaV	-0.19	ydcR	0.02	yfeA	0.08	sgbH	-0.09	otsA	0.11
ybaW	-0.01	ydcS	0.03	xapR	0.29	sgbE	-0.21	dcyD	-0.09
ybaW	-0.03	ydcS	0.00	xapB	0.05	yiaT	0.03	yedN	-0.06
ybaX	-0.01	ydcT	0.10	xapA	-0.03	yiaU	0.04	JW5315	-0.02
ybaE	0.03	ydcU	0.02	yfeN	-0.06	yiaV	0.02	yodD	0.08
cof	0.07	ydcU	-0.04	yfeR	0.02	yiaW	-0.03	yedS	0.08
ybaO	-0.08	ydcV	-0.02	yfeH	0.05	aldB	0.17	yedW	-0.01
mdlA	-0.06	ydcW	0.11	cysZ	-0.01	selB	0.35	yodB	0.08
glnK	-0.01	ydcX	0.01	cysK	-0.06	selA	0.17	yeeJ	0.18
amtB	0.03	ydcY	0.00	ptsH	-0.10	yibF	0.06	yeeL	0.06
tesB	-0.03	ydcZ	0.01	ptsI	0.09	rhsA	0.17	JW5326	0.02
ybaY	-0.08	yncB	-0.11	crr	1.01	yibA	-0.03	yeeP	-0.03
ybaZ	0.00	yncC	-0.04	pdxK	-0.03	yibG	0.10	dacD	0.17
ybaA	-0.27	yncD	-0.09	yfeS	-0.02	yibH	-0.13	nudD	0.08
ylaB	0.01	yncE	-0.02	cysM	0.02	yibI	0.06	yegH	0.12
ylaC	0.00	yncG	-0.02	cysA	0.19	yibI	0.04	yegM	0.08
maa	-0.09	yncH	0.04	cysW	0.04	mtlA	-0.03	yegP	0.00
hha	-0.04	rhsE	-0.05	cysU	-0.21	mtlD	0.31	gatR	0.14
ybaJ	-0.08	ydcD	-0.01	cysP	-0.27	mtlR	0.04	fbaB	0.11
acrB	0.01	JW5237	0.07	yfeT	-0.03	JW3576	-0.11	yegX	0.21
acrA	-0.01	ydcC	0.10	yfeT	0.04	yibL	0.19	yohN	0.15
acrR	0.01	ydcE	-0.03	yfeU	-0.02	lldP	-0.02	yehL	0.11
kefA	0.04	yddH	0.01	yfeV	-0.02	lldR	0.70	yehP	0.09
ybaM	0.03	yddH	0.05	yfeW	-0.02	lldD	0.05	yehR	0.05
priC	0.05	nhoA	0.00	yfeX	0.01	yibK	0.00	yehT	0.06
ybaN	0.02	yddE	0.05	ypeA	-0.05	cysE	-0.20	yehU	0.08
apt	-0.01	yddE	-0.01	ypeA	-0.12	secB	-0.09	JW5354	0.09
ybaB	0.04	narV	0.00	amiA	0.02	grxC	0.69	pbpG	0.08
recR	0.03	narW	-0.30	hemF	0.03	yibN	0.16	yohC	0.08
htpG	-0.05	narY	0.08	eutR	-0.01	gpmI	0.11	yeiW	0.06
aes	-0.06	narZ	0.04	eutK	-0.09	yibP	-0.01	yeiP	0.06
gsk	-0.20	narU	0.10	eutC	0.08	yibD	-0.08	yejO	0.08
ybaL	0.00	JW1465	-0.29	eutB	-0.04	tdh	0.01	yfaZ	-0.03
fsr	0.09	yddJ	-0.11	eutH	-0.27	kbl	-0.01	yfbE	0.11
ushA	0.13	yddK	-0.02	eutG	-0.02	rfaD	0.06	yfbJ	0.08
ybaK	-0.16	JW1468	-0.09	eutJ	0.00	rfaF	0.03	menD	0.11
ybaP	-0.03	yddG	-0.10	eutE	0.21	rfaC	-0.01	nuoC	-0.03
ybaQ	0.04	fdnG	-0.05	eutN	-0.31	rfaL	0.02	yfbT	0.08
copA	0.15	fdnH	0.04	eutM	-0.06	rfaZ	0.01	yfcE	0.17
ybaS	-0.02	fdnI	-0.01	eutD	0.06	rfaY	-0.02	yfcK	0.19
ybaT	0.07	yddM	0.01	eutP	0.01	rfaJ	-0.03	yfcM	0.04
cueR	0.05	adhP	0.02	eutS	0.09	rfaI	0.13	yfdI	0.05
ybbJ	-0.01	sfcA	-0.08	eutS	0.12	rfaB	0.02	tfaS	-0.05
ybbJ	0.12	rpsV	-0.20	ypfG	0.12	rfaS	0.07	yfdL	0.03
ybbK	-0.01	bdm	-0.04	ypfG	0.14	rfaP	-0.09	yfdN	0.01

ybbL	0.01	osmC	0.02	yffH	0.00	rfaG	0.03	JW5386	0.04
ybbM	-0.06	yddO	0.10	yffH	-0.04	rfaQ	-0.01	JW5387	0.11
ybbN	-0.04	yddP	0.05	aegA	0.03	mutM	-0.08	ypdA	0.05
ybbO	0.13	yddQ	0.01	narQ	-0.14	rpmG	-0.09	ypdH	0.16
tesA	-0.06	yddR	0.01	acrD	-0.17	pyrE	0.01	yfeA	0.12
ybbA	0.10	yddS	-0.12	yffB	-0.03	rph	0.08	JW5392	0.08
ybbP	-0.03	ddpX	0.08	JW2457	0.14	yicC	0.03	JW5393	0.04
rhsD	0.14	yddU	0.12	JW2457	-0.01	dinD	0.03	ucpA	0.02
ybbC	-0.03	yddV	-0.02	ypfH	0.08	yicG	0.22	yfeW	0.04
JW0488	0.01	yddW	-0.02	ypfH	0.08	yicF	-0.07	ypfH	-0.05
ybbD	0.09	xasA	0.03	ypfI	-0.01	rpoZ	-0.02	yfgE	0.02
ybbD	0.09	gadB	-0.14	ypfJ	0.02	trmH	-0.12	JW5398	0.05
JW5069	0.00	yddB	0.00	purC	0.02	recG	-0.05	yfgH	0.15
ybbB	0.06	ydeM	0.28	nlpB	0.01	gltS	-0.14	yfgJ	0.10
ybbS	0.03	ydeO	-0.02	gcvR	0.03	yicE	0.15	sseB	0.05
allA	0.08	ydeP	-0.03	bcp	0.04	yicH	-0.26	yphG	0.04
allR	0.37	ydeQ	-0.02	hyfA	0.04	yicI	0.08	yphH	0.05
gcl	0.04	ydeR	0.02	hyfB	-0.12	setC	-0.02	yfhK	0.08
hyi	-0.07	ydeS	-0.12	hyfC	0.05	yicL	0.25	yfhB	0.12
glxR	0.04	hipA	0.35	hyfD	0.01	nlpA	0.01	yfiP	0.03
ybbV	-0.02	hipB	-0.01	hyfE	-0.22	JW3636.1	-0.23	JW5410	0.09
ybbW	-0.03	ydeU	-0.01	hyfF	-0.13	yicP	0.03	JW5411	0.04
allB	0.02	ydeK	0.02	hyfG	-0.36	uhpT	-0.15	yfiL	0.04
ybbY	0.07	ydeV	-0.01	hyfH	0.12	uhpC	0.00	rimM	0.11
glxK	-0.03	ydeW	-0.03	hyfI	-0.10	uhpC	-0.04	yfjD	0.08
ylbA	-0.05	ego	0.05	hyfJ	-0.11	uhpB	-0.05	JW5417	0.12
allC	0.03	ydeY	0.00	focB	-0.22	uhpA	-0.03	yfjO	0.04
allD	0.12	ydeZ	0.03	perM	-0.12	ilvN	0.15	yfjP	-0.03
fdrA	-0.05	yneA	-0.05	yfgC	-0.02	ilvB	-0.02	ypjJ	0.05
ylbE	-0.02	yneB	0.02	yfgD	-0.07	ivbL	0.02	ypjA	0.06
ylbF	0.02	yneC	-0.02	uraA	-0.03	yidF	0.06	pinH	0.10
arcC	-0.07	tam	0.45	upp	0.04	yidG	-0.19	ypjC	0.09
purK	-0.09	yneE	-0.02	purM	-0.01	yidH	-0.05	ygaQ	0.04
purE	-0.02	uxaB	0.01	purN	-0.08	yidI	-0.10	JW5426	0.08
ppiB	-0.01	JW1515	0.02	ppk	0.01	yidJ	0.30	ygaT	0.08
ybcI	-0.02	JW1515	-0.05	ppx	0.14	yidK	-0.01	ygaY	0.08
sfmA	0.00	yneG	-0.05	yfgF	0.02	yidL	0.01	srlA	0.03
sfmC	-0.40	yneH	0.52	yfgF	0.01	JW3657	0.26	srlE	0.01
sfmD	-0.07	yneI	0.02	yfgH	-0.01	glvG	0.28	gutQ	0.08
fimZ	-0.06	yneJ	0.05	yfgH	0.03	glvB	0.02	ygaA	0.09
intD	0.11	yneK	-0.05	yfgI	0.03	glvC	0.17	ascG	-0.07
JW0526	0.10	sotB	0.23	guaA	0.03	yidP	0.04	JW5436	0.03
JW0527	-0.01	marC	-0.10	xseA	-0.09	yidE	0.04	rpoS	0.02
JW0530	0.01	marR	0.18	yfgJ	0.04	ibpB	-0.09	ygbF	-0.06
emrE	-0.27	marA	-0.12	yfgL	0.04	ibpA	-0.11	ygcI	0.05
ybcK	-0.02	marB	-0.05	yfgM	0.04	dgoK	-0.01	ygcQ	0.04
ybcL	-0.06	eamA	0.07	yfgA	0.07	yidA	0.00	ygcR	0.23
ybcM	0.00	eamA	-0.03	yfgA	0.03	yidB	-0.06	ygcU	0.06
ybcN	0.08	ydeE	0.03	yfgB	0.05	recF	-0.11	ygcW	-0.03
ninE	-0.02	ydeE	0.05	yfgB	0.09	trmE	0.02	ygcE	0.03
ybcO	0.01	ydeH	-0.03	ndk	0.00	tnaL	0.01	ygcG	0.06
rusA	-0.10	ydeI	-0.01	pbpC	0.00	tnaL	0.08	ygdI	0.04
ybcQ	0.07	ydeJ	0.06	yfhM	0.06	tnaA	0.02	amiC	0.07
JW0539.5	-0.01	dcp	-0.01	sseA	-0.04	yidY	-0.11	ygdB	0.07
essD	-0.06	ydfG	-0.11	sseB	-0.03	yidZ	-0.01	ppdB	-0.02
ybcS	-0.19	ydfH	-0.07	yfhJ	0.01	yieE	-0.19	JW5452	0.04
rzpD	-0.06	ydfZ	-0.06	fdx	0.07	yieF	0.09	yqeF	0.04
borD	-0.33	ydfI	-0.02	hscA	-0.08	yieG	-0.09	yqeH	-0.02
ybcV	0.47	ydfJ	-0.04	hscB	-0.82	yieH	-0.10	yqeJ	0.04
ybcW	0.03	pinQ	-0.01	yfhF	-0.01	yieI	0.10	ygeI	0.17
nohB	0.10	tfaQ	-0.02	nifU	0.05	yieJ	0.01	x	0.21
tfaD	0.03	tfaQ	0.00	iscS	0.00	yieC	0.06	ygeK	0.20
ybcY	0.01	stfQ	0.16	yfhP	-0.06	bglB	0.30	JW5459	0.07
ylcE	0.15	nohA	0.05	yfhQ	0.05	bglF	0.06	JW5460	0.15
appY	-0.12	gnsB	-0.02	yfhR	0.04	bglG	-0.04	ygeQ	0.21
ompT	-0.02	cspI	0.15	csiE	0.07	phoU	-0.01	xdhA	0.06

envY	-0.01	ydfP	-0.04	hcaT	0.04	pstB	-0.02	ygeW	0.02
ybcH	-0.04	ydfQ	0.06	hcaR	0.06	pstA	-0.05	guaD	0.07
nfrA	-0.08	ydfR	0.02	hcaR	0.16	pstC	0.18	ygfQ	0.00
nfrB	0.07	cspB	-0.03	hcaE	-0.03	pstS	-0.01	ygfS	0.04
cusS	-0.06	cspF	0.06	hcaF	-0.04	atpG	0.00	ygfT	0.01
cusR	0.07	ydfT	0.01	hcaC	0.02	atpA	-0.21	ygfU	0.04
cusC	-0.06	ydfU	0.07	hcaB	-0.13	atpH	-0.06	JW5471	0.04
cusF	-0.06	ydfU	0.03	hcaD	-0.09	atpF	-0.03	ygfB	0.04
cusA	0.05	rem	-0.07	yphA	-0.02	atpE	0.00	JW5474	0.09
pheP	-0.07	hokD	-0.07	yphB	0.15	atpB	-0.03	rpiA	0.10
ybdG	0.02	relE	-0.05	yphC	0.01	gidB	-0.05	ygfI	0.03
nfnB	0.15	relB	-0.10	yphD	-0.07	gidA	-0.02	yggP	0.07
ybdF	0.03	ydfV	0.01	yphE	0.10	mioC	-0.01	yggU	0.03
ybdJ	-0.01	flxA	0.00	yphF	0.07	asnC	-0.10	mltC	0.04
ybdK	0.10	ydfW	-0.01	yphF	-0.01	asnA	-0.01	yghE	0.00
hokE	-0.06	ydfX	0.04	yphG	0.10	yieN	-0.01	yghF	0.00
fepA	-0.04	dicC	0.06	yphG	-0.02	rbsD	-0.22	yghJ	0.08
fes	-0.10	ydfA	-0.01	yphH	0.01	rbsA	0.00	glcF	0.02
ybdZ	0.00	ydfC	0.01	glyA	0.01	rbsC	-0.10	glcE	0.02
entF	0.04	dicB	0.04	hmp	0.39	rbsB	0.09	yghO	0.10
fepE	-0.01	ydfD	0.08	glnB	0.02	rbsK	0.62	yghQ	0.07
fepC	0.00	ydfE	0.08	yfhA	-0.05	rbsR	0.21	yghS	0.00
fepG	-0.02	intQ	0.14	yfhG	0.00	yieO	0.01	yghU	0.05
fepD	-0.06	rspB	0.04	yfhK	-0.05	JW3736	0.28	JW5495	0.02
fepB	0.01	rspA	0.07	purL	0.19	yifE	-0.03	JW5496	0.10
entC	-0.03	ynfA	0.04	yfhD	0.04	yifB	-0.05	JW5497	0.13
entE	0.01	ynfB	0.05	yfhD	0.01	ilvL	-0.01	dkgA	0.05
entB	0.03	speG	0.01	yfhB	0.05	ilvL	-0.04	yqhG	0.00
entA	0.13	ynfC	0.02	yfhH	0.03	ilvG	0.02	ygiQ	0.09
ybdB	0.09	ynfE	0.16	yfhL	-0.09	ilvM	-0.01	tolC	0.13
ybdB	-0.01	ynfG	-0.06	pdxJ	-0.28	ilvE	0.00	ygiB	0.05
cstA	-0.04	mlc	-0.13	recO	0.00	ilvA	-0.02	yqiC	0.03
ybdD	0.06	ynfL	-0.03	lepA	0.07	ilvY	0.23	yqiG	0.05
ybdH	0.02	ynfM	0.02	rseC	-0.50	ilvC	0.05	yqiI	0.03
ybdL	0.03	asr	0.09	rseB	-0.12	ppiC	0.02	ygjO	-0.04
ybdM	-0.02	ydgD	0.04	rseA	-0.03	yifN	-0.06	ygjP	0.02
ybdN	-0.06	ydgE	0.00	nadB	0.05	rep	-0.07	ygjT	-0.08
ybdN	0.04	ydgF	0.24	yfiC	0.06	rhlB	-0.01	yqjC	0.00
ybdO	0.17	ydgG	-0.05	srmB	-0.03	rhoL	-0.04	yhaL	-0.03
dsbG	0.12	ydgG	0.06	yfiE	0.09	rhoL	0.14	yhaM	0.03
ahpC	-0.07	pntB	0.02	yfiK	0.08	rho	-0.30	yhaO	0.01
ahpF	0.07	pntA	0.08	yfiD	0.04	wecA	-0.07	tdcG	0.03
ybdQ	-0.22	ydgH	0.02	ung	0.02	wecB	0.09	tdcE	0.00
ybdR	-0.10	ydgH	-0.02	yfiF	0.00	wecC	0.06	tdcR	0.05
rnk	-0.10	ydgI	0.02	trxC	0.06	rffG	-0.09	garR	-0.03
rna	0.04	ydgB	-0.03	yfiP	0.06	rffH	0.00	yhbO	0.18
citT	-0.12	ydgC	0.03	yfiQ	0.03	rffC	-0.04	deaD	0.19
citG	-0.06	rstA	0.06	yfiM	0.01	rffA	-0.02	pnp	0.15
citX	-0.25	rstB	0.07	kgtP	-0.05	wzxE	0.00	yhbC	-0.01
citE	0.16	tus	0.06	clpB	0.15	wecG	0.06	yhbX	0.00
citD	-0.08	fumC	-0.04	JW2574	0.00	yifK	0.04	nanK	0.00
citC	0.06	fumA	-0.09	yfiH	0.07	aslB	0.00	yhcB	-0.01
dpiB	0.06	manA	-0.05	yfiH	0.13	aslA	-0.16	yhcN	-0.02
dpiA	0.09	ydgA	-0.01	rluD	0.03	hemY	0.05	yhcR	-0.05
dcuC	-0.05	uidC	0.63	yfiA	0.00	hemX	0.01	yhdP	0.16
crcA	-0.09	uidB	0.04	pheL	-0.06	cyaA	0.02	yhdJ	-0.01
cspE	-0.02	uidA	-0.01	pheA	-0.06	cyaY	0.09	zraP	0.03
crcB	-0.03	uidR	0.19	tyrA	-0.11	yzcX	-0.02	nfi	0.04
ybeM	-0.02	hdhA	-0.05	aroF	0.05	yifL	-0.15	thiG	-0.01
tatE	0.10	malI	-0.08	yfiL	0.06	yigA	-0.04	sthA	0.01
lipA	-0.03	malY	0.13	yfiR	0.05	xerC	-0.06	JW5552	0.03
ybeF	-0.04	add	0.45	yfiN	0.01	yigB	0.09	argB	-0.02
ybeD	-0.06	ydgJ	-0.09	yfiB	-0.02	uvrD	-0.23	JW5554	-0.01
dacA	-0.05	blr	0.04	ypjD	0.05	corA	-0.07	gldA	-0.06
rlpA	-0.05	ydgT	-0.01	yfjD	-0.11	yigF	0.13	cpxP	0.10
ybeA	0.07	ydgK	-0.01	recN	0.06	pldA	0.07	yiiM	-0.02

ybeB	0.01	rnfA	0.08	smpA	-0.39	JW3795	-0.14	yiiF	-0.01
cobC	0.02	rnfB	0.17	yfjF	0.04	pldB	0.02	yihS	0.01
ybeL	-0.03	rnfC	-0.03	yfjG	0.06	yigM	0.02	yihO	0.10
ybeQ	-0.07	rnfD	0.14	smpB	-0.05	metR	0.30	typA	0.08
ybeR	-0.05	rnfG	-0.03	intA	0.03	metE	0.01	JW5572	0.10
ybeS	0.00	rnfE	0.11	yfjH	0.07	udp	-0.04	yihF	0.07
ybeT	-0.06	nth	0.30	alpA	0.06	udp	-0.01	mobB	0.00
ybeU	0.04	ydgR	0.02	yfjI	0.07	rmuC	0.47	trkH	0.10
ybeV	0.00	gst	-0.37	yfjI	0.01	ubiE	-0.07	yigZ	0.18
hscC	-0.26	pdxY	-0.03	JW2606	0.00	tatA	0.00	tatD	-0.02
rihA	0.52	pdxH	0.03	yfjJ	0.09	tatC	-0.11	tatB	0.08
gltL	-0.06	ydhA	0.06	yfjK	0.08	fre	0.01	yigL	0.04
gltK	-0.04	ydhH	0.08	yfjL	-0.01	fadA	0.01	rhtB	0.06
gltJ	0.03	slyB	-0.04	yfjM	0.06	fadB	0.06	rhtC	0.03
JW0651	0.08	slyA	0.21	yfjN	0.05	pepQ	0.05	yigI	0.04
corC	0.08	ydhI	-0.06	yfjO	0.00	yigZ	-0.01	rarD	0.08
ybeY	0.02	ydhJ	-0.02	yfjP	0.09	mobA	0.06	yigG	0.11
ybeZ	0.01	ydhJ	0.10	yfjQ	0.03	yihD	-0.09	yigE	0.00
yleA	-0.06	ydhK	0.06	yfjR	-0.01	yihE	-0.03	dapF	0.12
ubiF	-0.02	sodC	0.33	JW2616	0.00	dsbA	0.15	wecF	0.03
asnB	-0.06	ydhF	-0.05	yfjS	0.09	yihG	-0.08	wzzE	0.01
nagD	-0.01	JW1640	0.07	yfjT	-0.31	polA	0.45	trxA	0.01
nagC	-0.10	ydhM	-0.04	yfjU	0.06	yihI	-0.05	gppA	0.01
nagA	-0.03	nemA	0.14	yfjV	-0.04	hemN	0.15	ilvD	0.04
nagB	-0.12	gloA	-0.01	yfjW	-0.01	glnG	-0.06	hdfR	0.03
nagE	0.24	rnt	0.07	JW2623.5	0.10	glnL	-0.01	yieP	0.00
ybfM	-0.09	lhr	0.01	yfjX	0.09	glnA	-0.06	kup	-0.05
ybfN	0.03	ydhD	-0.07	yfjY	0.02	yihL	-0.01	yieM	0.08
fur	-0.08	ydhD	0.12	yfjZ	0.00	yihM	-0.03	atpI	0.09
ybfJ	0.08	ydhO	0.04	ypjF	-0.01	yihN	-0.02	yieL	0.00
ybfE	-0.04	ydhO	-0.26	ypjA	-0.03	yshA	0.15	yieK	0.12
ybfF	-0.09	sodB	-0.02	ypjB	-0.03	yihP	0.09	tnaB	0.01
seqA	-0.29	ydhP	-0.06	ypjC	0.01	yihQ	0.00	yidX	0.05
pgm	-0.03	purR	-0.02	JW2631	0.02	yihR	-0.08	dgoR	0.07
ybfP	-0.06	ydhB	0.10	JW5426	-0.01	yihT	-0.03	dgoA	0.02
ybfG	0.03	ydhC	0.04	yqaD	-0.03	yihU	0.00	dgoT	0.05
potE	0.14	ydhC	-0.15	ygaT	0.05	rbn	-0.05	yidS	0.01
speF	-0.13	cfa	-0.04	ygaF	0.06	dtd	-0.19	yidR	-0.04
ybfK	-0.08	norM	-0.02	ygaF	0.04	yiiD	-0.01	yidR	0.01
ybfK	-0.01	norM	0.11	gabD	0.03	fdhE	0.56	yidQ	0.04
kdpE	0.05	ydhQ	-0.03	gabT	0.09	fdoI	0.09	emrD	0.01
kdpD	0.06	ydhR	0.03	gabP	-0.31	fdoH	0.02	JW5635	0.09
kdpC	0.11	ydhS	0.04	ygaE	0.00	fdoG	0.22	yicO	0.07
kdpB	0.01	ydhS	0.00	ygaU	0.12	fdhD	0.01	yicN	0.00
kdpA	0.04	ydhT	0.02	yqaE	0.05	yiiG	-0.02	yicM	-0.05
kdpF	0.17	ydhU	0.02	ygaV	0.03	frvR	0.23	JW5639	-0.03
ybfA	0.01	ydhW	0.00	ygaP	0.55	frvX	0.02	ttk	0.03
rhsC	0.15	ydhV	-0.02	stpA	0.19	frvA	0.01	radC	0.02
ybfB	0.06	ydhY	0.00	ygaW	0.06	rhaS	0.06	htrL	0.04
ybfB	0.01	ydhY	-0.01	ygaC	0.05	rhaR	-0.12	yibQ	0.02
ybfO	0.21	ydhZ	-0.03	ygaM	0.03	rhaT	0.17	yibP	0.08
ybfC	-0.05	pykF	-0.07	nrdH	0.36	sodA	-0.16	yiaY	0.02
JW0694	-0.05	lpp	-0.16	nrdI	0.06	cpxA	-0.01	JW5649	0.04
ybfL	-0.05	ynhG	-0.02	nrdE	-0.01	cpxR	0.00	sgbU	0.06
ybfD	0.07	sufE	0.00	nrdF	0.01	yiiP	0.10	yiaN	0.02
ybgA	-0.02	csdB	0.06	proV	-0.32	pfkA	0.17	yiaB	0.03
phrB	0.39	sufD	0.10	proW	-0.14	sbp	0.46	yiaF	0.03
ybgH	0.04	sufC	0.10	proX	0.04	cdh	-0.20	tkrA	0.03
ybgI	-0.03	sufB	-0.05	ygaY	0.02	tpiA	0.14	bisC	0.04
ybgJ	0.06	sufB	0.03	ygaZ	-0.02	yiiQ	-0.39	yhjY	0.00
ybgK	-0.10	sufA	0.09	ygaH	0.08	yiiR	0.42	yhjW	0.10
ybgL	-0.08	ydiH	-0.01	mprA	-0.16	yiiS	-0.09	yzpK	0.07
nei	-0.38	ydiI	-0.04	emrA	0.04	yiiT	0.11	JW5662	-0.01
abrB	-0.14	ydiJ	-0.18	emrB	0.02	fpr	-0.10	yhjT	0.03
ybgO	-0.02	ydiK	-0.03	luxS	-0.04	glpX	-0.02	yhjQ	0.05
ybgP	0.06	ydiM	-0.04	gshA	0.02	glpK	-0.07	bcsA	0.05

ybgQ	-0.05	ydiM	0.01	yqaA	-0.01	glpF	0.39	bcsC	0.04
ybgD	-0.02	ydiN	-0.07	yqaB	0.00	yiiU	-0.07	yhjK	0.03
gltA	0.03	ydiN	-0.02	yqaB	-0.31	menG	0.03	kdgK	0.09
sdhC	-0.13	ydiB	0.02	alaS	-0.02	menA	-0.01	hdeB	0.03
sdhD	-0.07	ydiB	-0.02	recX	0.03	cytR	-0.11	yhiD	-0.01
sdhA	-0.05	aroD	0.07	ygaD	0.02	priA	0.01	yhiQ	0.00
sdhB	-0.08	ydiF	0.01	mltB	0.00	rpmE	-0.11	yhiM	0.04
sucA	-0.09	ydiF	-0.02	srlB	-0.23	yiiX	-0.02	JW5675	0.04
sucB	-0.09	ydiO	0.01	srlD	-0.17	metJ	0.19	yhiH	0.04
sucC	-0.07	ydiP	-0.07	gutM	0.02	metB	-0.01	yhhJ	0.02
sucD	0.08	ydiP	0.02	srlR	0.16	metL	-0.21	rhsB	-0.06
farR	0.11	ydiQ	-0.01	gutQ	-0.04	metF	0.01	yhhT	0.03
hrsA	0.38	ydiR	-0.07	JW2678	-0.04	katG	0.19	yhhS	0.00
ybgG	-0.04	ydiS	-0.02	JW2679	-0.04	yijF	-0.04	dcrB	0.00
ybgG	0.00	ydiT	0.04	flrD	-0.07	fsaB	-0.06	yhhL	0.00
cydB	0.00	ydiD	0.02	flrR	0.01	frwC	0.02	yrhD	0.04
ybgT	-0.02	ydiD	0.15	hypF	0.30	frwB	0.02	gntR	0.01
ybgE	-0.08	ppsA	-0.07	hydN	-0.12	pflD	-0.04	gntU	0.09
ybgC	0.02	ydiA	-0.05	ascF	-0.01	pflC	0.15	rtcA	-0.01
ybgC	-0.04	aroH	0.06	ascB	0.06	frwD	0.00	yhgH	0.01
tolQ	0.00	ydiE	0.05	hycG	-0.12	yijO	-0.02	hslO	0.03
tolR	-0.44	ydiU	0.04	hycF	-0.24	yijP	-0.03	yrfD	0.02
tolA	-0.07	ydiV	0.06	hycE	-0.23	ppc	0.11	yrfA	0.03
pal	-0.01	nlpC	-0.02	hycD	0.01	argE	0.23	aroK	0.08
ybgF	0.06	btuD	-0.29	hycC	-0.04	argC	0.04	yhfY	-0.01
nadA	0.04	btuE	-0.57	hycB	-0.01	argH	0.30	yhfU	0.02
pnuC	0.06	btuC	-0.23	hycA	-0.03	oxyR	0.07	yhfR	0.06
zitB	-0.06	ihfA	0.17	hypA	-0.29	yijC	-0.06	yhfO	0.04
ybgS	-0.07	pheM	-0.27	hypB	0.06	yijD	-0.01	yhfN	0.04
aroG	-0.17	rpmI	-0.15	hypC	-0.04	trmA	-0.08	yhfK	0.05
galM	-0.06	arpB	0.06	hypD	0.02	btuB	0.56	yheO	0.02
galK	-0.27	ydiY	0.02	hypE	-0.01	coaA	-0.06	pshM	0.06
galT	-0.22	pfkB	0.17	fhlA	0.01	tufB	-0.06	gspL	0.08
galE	-0.71	ydiZ	0.04	ygbA	0.10	rplK	0.04	gspI	0.02
modF	-0.08	yniA	0.09	mutS	0.08	rplA	-0.06	gspD	0.05
modE	0.69	yniB	0.05	pphB	-0.02	htrC	-0.02	smf	0.01
ybhT	-0.07	yniC	-0.01	ygbI	-0.04	thiH	0.09	yrdD	0.02
modA	0.41	yniC	0.02	ygbJ	-0.03	thiS	-0.10	yrdA	-0.02
modB	0.39	ydjM	-0.01	ygbJ	-0.06	thiF	0.02	yjbF	0.03
modC	0.34	ydjN	-0.07	ygbK	-0.03	thiE	0.05	JW5712	0.02
ybhA	0.00	ydjO	-0.22	ygbL	0.05	thiC	0.03	ubiC	-0.05
ybhE	0.07	cedA	0.05	ygbM	0.04	rsd	0.02	zur	0.00
ybhD	-0.02	katE	0.11	ygbN	0.02	nudC	0.06	yjbN	-0.05
ybhH	0.07	ydjC	-0.04	rpoS	-0.10	hemE	-0.11	yjbO	0.09
ybhI	0.01	ydjC	0.00	pcm	-0.01	yjaG	0.08	yjbS	0.03
ybhJ	0.04	celF	-0.45	surE	-0.01	hupA	-0.15	yjcB	0.03
ybhC	-0.06	celD	-0.11	ygbO	0.01	yjaH	-0.17	JW5719	0.01
ybhB	0.05	celC	-0.10	ygbE	0.04	zraS	-0.02	JW5720	0.01
bioA	-0.10	celB	-0.08	cysC	-0.04	zraR	-0.03	yjcS	0.00
bioB	0.09	celA	-0.03	cysN	0.01	purD	0.01	phnQ	0.04
bioF	0.05	osmE	0.07	cysD	-0.02	purH	-0.02	phnQ	-0.02
bioC	0.17	cho	0.07	iap	0.03	yjaA	0.02	JW5728	-0.01
bioD	0.11	spy	-0.09	ygbF	0.03	yjaB	0.03	yjcZ	0.06
uvrB	0.12	astE	0.06	ygbT	0.07	metA	0.12	yjdB	0.03
ybhK	0.04	astB	0.06	ygcH	0.07	aceB	0.15	adiA	0.08
moaA	0.19	astA	0.04	JW2727	0.04	aceA	0.15	yjdO	0.02
moaB	0.04	argM	0.02	ygcJ	-0.03	aceK	-0.05	yjdC	0.01
moaC	0.37	xthA	0.10	ygcK	0.10	arpA	0.18	dsbD	0.10
moaD	0.11	ydjX	-0.02	ygcL	0.00	iclR	0.30	dcuA	0.11
moaE	-0.30	ydjZ	0.00	ygcB	0.05	metH	0.14	yjeI	0.07
ybhL	0.57	ynjA	-0.17	cysH	-0.12	yjbB	0.14	ecnA	0.01
ybhM	-0.01	ynjB	0.09	cysI	-0.06	pepE	0.01	sugE	0.03
ybhM	0.13	ynjE	-0.06	cysJ	0.07	yjbC	-0.09	yjeM	0.07
ybhN	-0.01	ynjF	0.06	ygcM	-0.01	yjbD	-0.18	rnr	0.06
ybhO	-0.33	ynjF	-0.03	ygcN	-0.01	lysC	0.21	yjfN	-0.01
ybhP	0.00	nudG	0.05	ygcO	0.01	pgi	0.14	yjfO	0.03

ybhP	0.03	ynjH	0.01	ygcP	-0.10	yjbE	0.00	sgaT	0.03
ybhQ	-0.07	gdhA	0.07	ygcQ	0.11	yjbG	-0.06	ytfB	-0.02
ybhR	-0.17	ynjI	0.04	ygcR	0.01	yjbH	0.27	fklB	-0.01
JW0776	0.01	topB	-0.02	ygcS	-0.12	psiE	0.32	ytfH	0.05
ybhS	-0.03	selD	0.21	ygcU	-0.01	xylE	0.08	ytfI	0.01
ybhG	-0.06	ydjA	0.05	ygcU	0.05	malG	0.22	ytfK	0.01
ybiH	0.15	ansA	-0.12	JW2743	0.05	malF	0.29	JW5751	0.08
rhlE	-0.10	pncA	-0.02	ygcW	0.05	malE	-0.10	ytfR	0.08
JW0782	0.00	pncA	-0.46	yqcE	0.00	malK	0.00	ytfT	0.06
ybiA	0.04	ydjE	-0.12	ygcE	0.03	lamB	0.02	yjfF	0.03
dinG	0.04	ydjF	0.15	ygcF	0.00	malM	-0.04	yjgF	-0.03
ybiB	-0.02	ydjG	-0.14	ygcG	0.03	malM	0.09	yjgK	0.01
ybiC	-0.15	ydjH	0.00	mazG	0.01	yjbI	-0.04	yjgL	0.01
ybiJ	-0.07	ydjH	0.01	chpA	0.06	dgkA	0.15	yjgM	0.02
ybiI	-0.13	ydjI	0.04	relA	-0.01	dinF	-0.08	yjgN	-0.01
ybiX	-0.01	ydjJ	-0.10	rumA	-0.03	yjbJ	-0.03	yjgB	-0.02
ybiL	-0.03	ydjL	0.38	barA	-0.02	yjbL	-0.05	JW5765	0.06
ybiN	-0.04	yeaC	0.05	gudD	0.43	yjbM	-0.04	JW5766	0.03
ybiO	0.00	msrB	0.03	gudX	-0.02	qor	-0.13	yjhB	0.02
glnQ	0.54	yeaD	-0.08	gudP	0.02	alr	0.23	yjhC	0.01
glnP	0.84	yeaE	-0.09	yqcA	-0.08	tyrB	0.02	yjhD	0.04
glnH	0.08	mipA	0.06	yqcB	0.02	aphA	-0.01	yjhU	0.02
dps	0.04	yeaG	-0.02	yqcC	0.01	JW4016	0.01	yjhH	0.10
ybiF	0.06	yeaH	-0.07	syd	0.06	yjbQ	-0.14	sgcX	0.09
ompX	0.01	yeaI	-0.02	yqcD	0.07	yjbR	-0.14	yjhT	-0.02
ybiP	-0.03	yeaI	0.04	yqcD	-0.06	uvrA	-0.03	yjhA	0.11
mntR	-0.02	yeaJ	0.05	ygdH	-0.11	yjcC	0.13	fimI	0.04
ybiR	-0.04	yeaJ	0.02	sdaC	0.18	soxS	0.23	JW5781	0.07
ybiS	0.02	yeaK	0.04	sdaB	-0.12	soxR	0.27	yjiD	0.06
ybiT	0.03	JW1777	-0.44	fucO	0.06	yjcD	0.12	yjiH	0.04
ybiU	-0.08	yeaL	0.07	fucA	0.01	yjcE	-0.07	kptA	0.04
ybiV	0.16	yeaM	-0.13	fucP	-0.05	yjcF	-0.07	kptA	-0.03
ybiW	-0.04	yeaN	0.14	fucI	0.10	yjcG	-0.13	yjiL	-0.01
ybiY	-0.12	yeaO	0.09	fucK	-0.03	yjcH	-0.05	yjiM	0.08
fsaA	-0.07	yoaF	0.03	fucU	0.00	acs	0.41	yjiT	0.05
moeB	-0.03	yeaQ	-0.05	fucR	0.27	nrfA	-0.02	mcrD	0.10
moeA	0.19	yoaG	0.01	ygdE	0.02	nrfB	0.05	mcrC	0.04
ybiK	0.06	yeaR	0.01	ygdD	0.07	nrfC	0.15	yjiA	0.01
yliC	-0.04	yeaR	-0.03	gcvA	0.07	nrfD	0.82	yjiY	0.08
yliD	0.01	yeaS	0.05	ygdI	-0.06	nrfE	0.27	yjjM	0.02
yliE	-0.11	yeaS	0.16	csdA	-0.03	nrfF	0.55	yjjN	-0.01
yliF	-0.14	yeaT	0.05	ygdK	-0.02	nrfG	0.31	mdoB	0.04
yliG	-0.08	yeaU	0.08	ygdL	-0.16	gltP	0.04	yjjA	0.04
yliH	0.08	yeaX	0.02	mltA	0.01	yjcO	-0.06	yjjP	0.08
yliI	0.03	rnd	0.10	argA	-0.03	fdhF	0.43	yjjZ	0.03
yliJ	-0.08	fadD	0.19	recD	-0.07	yjcP	-0.26	prfC	0.01
dacC	-0.04	yeaY	0.01	recB	0.02	yjcQ	0.25	JW5799	0.04
deoR	0.23	yoaA	-0.04	ptrA	0.04	yjcR	-0.12	nadR	0.18
ybjG	0.11	yoaB	0.03	recC	0.00	alsE	0.04	yjjX	0.04
cmr	0.00	yoaC	-0.25	ppdC	0.02	alsC	0.14	ybhR	0.12
ybjH	-0.10	yoaH	0.10	ygdB	-0.08	alsA	-0.03	hyfI	0.18
ybjI	-0.04	pabB	0.07	ppdB	0.01	alsB	0.18	leuB	-0.02
ybjJ	-0.03	yeaB	0.05	ppdA	0.10	rpiR	0.10	pcnB	0.03
ybjK	0.02	yeaB	-0.01	thyA	-0.12	rpiB	0.15	yadB	0.01
ybjL	0.02	sdaA	-0.04	thyA	-0.07	phnP	-0.07	fhiA	0.00
ybjM	0.17	yoaD	-0.04	ptsP	-0.07	phnO	-0.05	mbhA	0.05
grxA	0.12	yoaE	0.00	nudH	0.00	phnN	-0.05	ykfC	0.06
ybjC	-0.08	yoaE	0.18	nudH	-0.04	phnM	0.20	JW5814	0.03
nfsA	-0.01	manX	0.04	mutH	0.07	phnL	-0.01	tfaD	-0.03
rimK	-0.12	manY	0.02	ygdQ	-0.01	phnK	0.72	ybfE	-0.06
ybjN	0.00	manZ	-0.25	ygdQ	-0.05	phnJ	-0.07	ybhD	0.02
potF	0.25	yobD	0.00	ygdR	0.05	phnI	-0.17	potG	0.03
potG	-0.05	JW1810	0.04	tas	0.69	phnH	0.09	ybjS	0.00
potH	-0.01	rrmA	0.00	ygeD	-0.02	phnG	-0.01	ycfS	-0.01
potI	-0.07	cspC	0.01	ygeD	0.00	phnF	0.24	mltE	-0.04
ybjO	0.02	yobF	-0.02	aas	0.15	phnE	-0.21	abgT	0.04

ybjF	0.03	yebO	0.02	galR	0.28	phnD	-0.01	ydcH	-0.03
artJ	0.08	yobG	-0.35	lysA	-0.14	phnC	0.24	yddM	-0.03
artM	-0.02	kdgR	0.03	lysR	0.22	phnB	-0.05	yneF	0.01
artQ	-0.01	yebR	0.06	ygeA	0.03	phnA	0.15	asr	0.08
artI	0.06	yebS	0.15	araE	0.03	yjdA	-0.06	ydhL	0.03
artP	0.01	yebT	-0.03	kduD	-0.01	proP	0.02	ydiD	0.05
ybjP	-0.01	yebU	0.05	kduI	0.33	basS	-0.06	yebN	0.02
ybjQ	-0.01	yebV	0.00	yqeG	-0.41	basR	0.02	znuA	0.00
ybjQ	-0.07	yebW	0.01	yqeH	0.04	yjdE	-0.04	yedQ	0.00
ybjR	0.05	pphA	0.19	yqeI	0.04	adiY	0.25	yeeJ	-0.05
ybjS	-0.02	yebY	0.00	yqeJ	-0.02	melR	0.26	yeeY	0.02
ybjT	-0.02	yebZ	-0.02	yqeK	0.10	melA	-0.02	wzzB	0.01
ltaE	0.08	yobA	-0.08	ygeF	-0.04	melB	0.39	yegR	0.01
poxB	0.01	holE	0.20	ygeG	0.10	yjdF	-0.05	yohG	0.01
hcp	0.08	yobB	0.07	ygeH	0.13	fumB	-0.09	yejO	0.01
ybjE	0.06	yobB	-0.01	ygeI	-0.01	dcuB	0.28	elaD	0.00
aqpZ	0.00	exoX	0.04	x	0.00	dcuR	0.00	yfcB	-0.03
ybjD	-0.12	yebE	0.01	ygeK	-0.04	dcuS	0.07	yphC	-0.01
ybjX	0.07	yebF	-0.06	JW5460	0.04	yjdI	-0.06	ygaA	0.04
macA	0.07	yebG	-0.03	ygeP	0.07	yjdJ	-0.01	ygcI	0.00
macB	0.01	purT	0.06	ygeQ	0.03	yjdK	-0.07	ygcS	0.00
cspD	-0.12	eda	0.02	ygeR	-0.01	lysU	0.11	JW5846	0.01
yljA	0.08	edd	-0.04	xdhB	0.33	yjdL	0.21	prfB	-0.02
clpA	-0.02	zwf	0.07	xdhC	0.01	cadA	-0.05	yghO	0.00
aat	0.06	hexR	0.10	ygeV	0.04	cadB	-0.18	yqhC	-0.03
cydD	-0.09	pykA	0.15	ygeX	0.02	cadC	-0.05	yqjF	-0.07
trxB	-0.02	msbB	0.04	ygeY	0.06	cutA	-0.02	pnp	-0.03
lrp	0.23	yebA	0.35	ygeZ	-0.27	aspA	-0.05	yihO	-0.03
ycaJ	-0.01	znuA	-0.12	yqeA	-0.09	fxsA	0.58	ysgA	-0.04
dmsB	0.02	znuC	-0.43	yqeB	0.00	yjeH	0.01	yigL	-0.02
dmsC	-0.20	znuB	-0.07	ygfJ	0.02	groL	-0.06	recQ	0.00
ycaC	-0.02	ruvB	-0.02	ssnA	0.02	yjeJ	-0.08	trxA	-0.03
ycaD	-0.17	ruvA	0.02	ygfM	0.06	yjeK	-0.13	rbsD	-0.06
ycaM	0.27	yebB	0.04	xdhD	-0.07	efp	-0.07	yidX	-0.01
ycaN	-0.08	ruvC	0.02	ygfO	-0.03	ecnB	0.56	dgoT	-0.03
ycaK	-0.09	yebC	0.08	idi	0.02	blc	-0.06	yidR	-0.03
pflA	-0.04	nudB	0.02	lysS	0.11	ampC	-0.01	yicJ	0.00
pflB	-0.06	yecD	0.23	recJ	-0.08	frdD	0.00	bcsC	-0.04
focA	0.12	yecE	-0.04	dsbC	0.34	frdC	0.09	yhjK	-0.02
ycaO	-0.02	yecN	-0.01	xerD	0.09	frdB	0.13	yrhA	-0.03
ycaO	-0.01	yecO	0.32	fldB	0.06	frdA	0.14	aroK	-0.05
ycaP	0.06	yecP	0.29	ygfX	-0.08	poxA	-0.01	aidB	-0.02
serC	0.04	torZ	-0.29	ygfY	-0.05	yjeN	-0.35	yjfR	-0.04
aroA	0.03	torY	-0.01	ygfZ	-0.08	yjeO	-0.44	yjiK	-0.03
ycaL	0.16	cutC	-0.10	yqfA	-0.08	yjeP	-0.66	mcrB	-0.04
cmk	-0.09	yecM	-0.02	yqfB	-0.06	yjeQ	0.16	bglJ	-0.03
ihfB	0.22	yecT	0.13	bglA	-0.08	yjeS	-0.04	prfC	-0.05
ycaI	-0.02	flhE	0.00	ygfF	-0.02	amiB	0.17	ydhM	-0.03
ycaQ	-0.03	flhA	0.10	gcvP	-0.06	mutL	0.04	yfeH	-0.05
ycaR	0.08	flhB	-0.07	gcvH	0.00	miaA	-0.17	ypeB	-0.03
ycbJ	-0.10	cheZ	-0.11	gcvT	-0.09	hfq	-0.06	csiE	-0.02
ycbC	0.05	cheY	-0.08	visC	0.11	hflC	0.02	JW5710	-0.02
smtA	0.10	cheB	-0.06	ubiH	0.06	yjeT	0.16	JW5711	0.03
ycbB	0.06	cheR	-0.26	pepP	-0.45	purA	-0.01	JW5712	0.02
ycbK	0.03	tap	0.01	ygfE	0.06	yjeB	-0.05	JW5713	-0.05
ycbL	-0.03	tar	0.05	ygfE	-0.06	yjfH	-0.04	JW5714	0.00
aspC	0.04	cheW	0.44	ygfA	0.00	yjfI	0.46	JW5715	-0.05
ompF	-0.03	cheA	0.02	serA	0.02	yjfJ	0.34	JW5716	0.09
pncB	0.08	motB	0.02	rpiA	-0.02	yjfK	0.55	JW5717	0.03
pepN	-0.03	motA	0.05	yqfE	-0.01	yjfL	0.55	JW5718	0.03
ssuB	0.01	flhC	0.12	iciA	0.26	yjfM	0.35	JW5719	0.01
ssuD	-0.02	flhD	0.02	sbm	0.03	yjfC	0.20	JW5720	0.01
ssuA	0.05	yecG	0.04	argK	0.04	aidB	-0.07	JW5721	0.00
ssuE	0.00	otsA	-0.01	ygfG	0.11	yjfP	0.12	JW5725	0.04
ycbR	-0.08	otsB	-0.01	ygfH	-0.04	yjfQ	0.01	JW5726	-0.02
ycbS	-0.06	araH	-0.07	ygfI	-0.01	yjfR	-0.08	JW5728	-0.01

ycbT	-0.09	araG	0.07	yggE	-0.05	sgaB	-0.02	JW5729	0.06
ycbU	-0.11	araF	-0.15	yggA	-0.05	sgaA	-0.05	JW5730	0.03
ycbV	-0.09	ftnB	0.07	yggB	0.53	sgaH	0.13	JW5731	0.08
ycbF	0.32	ftnB	0.05	epd	0.03	sgaU	-0.01	JW5732	0.02
pyrD	0.01	yecJ	-0.06	yggC	0.03	sgaE	0.06	JW5733	0.01
ycbW	-0.12	yecR	0.04	yggD	-0.02	yjfY	0.08	JW5734	0.10
ycbX	-0.05	yecR	0.09	yggF	-0.04	rpsF	-0.11	JW5735	0.11
ycbY	0.05	ftnA	-0.36	cmtA	-0.06	priB	0.11	JW5736	0.07
uup	-0.02	yecH	0.07	cmtB	-0.02	rplI	-0.12	JW5737	0.01
pqiA	0.01	tyrP	0.08	tktA	-0.05	yjfZ	0.00	JW5738	0.03
pqiB	-0.09	yecA	-0.02	yggG	0.03	ytfA	-0.01	JW5739	0.07
ymbA	0.02	uvrC	-0.25	speB	0.02	cycA	0.37	JW5741	0.06
rmf	0.15	uvrY	0.12	speA	-0.05	ytfE	0.06	JW5742	-0.01
ycbZ	0.07	yecF	0.00	yqgB	-0.06	ytfF	0.41	JW5743	0.03
ycbG	-0.01	sdiA	0.19	yqgC	0.03	ytfG	-0.02	JW5744	0.03
ompA	0.01	yecC	0.02	galP	0.06	cpdB	0.03	JW5745	-0.02
sulA	0.02	yecS	0.01	sprT	0.02	cysQ	-0.02	JW5746	-0.01
yccR	-0.03	fliY	-0.06	endA	0.07	ytfJ	0.28	JW5747	0.05
yccS	0.10	fliZ	0.00	yggJ	-0.24	ytfL	-0.06	JW5748	0.01
yccF	0.04	fliA	-0.01	gshB	-0.01	msrA	-0.09	JW5749	0.01
helD	0.02	fliC	-0.08	yqgE	-0.03	ytfM	0.04	JW5751	0.08
yccT	0.01	fliD	-0.02	yggR	-0.14	ytfN	0.16	JW5752	0.08
yccU	0.00	fliS	0.07	yggS	0.05	ytfP	0.07	JW5753	0.06
yccV	-0.03	fliT	0.25	yggT	-0.07	yzfA	0.03	JW5754	0.03
yccW	-0.10	amyA	0.01	yggV	-0.06	chpB	0.36	JW5755	-0.03
yccX	-0.04	yedD	0.09	yggW	0.02	ytfQ	0.11	JW5756	0.01
yccK	0.04	yedE	0.12	yggM	-0.02	fbp	0.01	JW5757	0.01
yccA	-0.14	yedF	0.07	ansB	0.08	mpl	0.01	JW5758	0.02
hyaA	-0.07	yedK	-0.07	yggN	-0.02	yjgA	-0.02	JW5759	-0.01
hyaB	-0.03	yedL	0.04	yggL	0.02	pmbA	-0.07	JW5761	-0.02
hyaC	-0.49	yedN	0.05	yggH	-0.04	cybC	0.01	JW5763	0.04
hyaD	-0.01	yedN	0.08	mutY	-0.03	nrdG	-0.03	JW5764	0.09
hyaE	-0.11	yedM	0.38	yggX	0.00	nrdD	0.05	JW5765	0.06
hyaF	-0.10	fliE	-0.10	mltC	0.04	treC	0.32	JW5766	0.03
appC	-0.10	fliF	0.08	JW2931	-0.03	treB	0.14	JW5768	0.02
appB	0.05	fliG	0.03	nupG	-0.13	treR	0.28	JW5769	0.01
yccB	0.05	fliH	0.06	yqgA	-0.03	mgtA	0.18	JW5770	0.04
appA	0.01	fliI	0.04	yghD	-0.20	pyrI	0.39	JW5774	0.02
etk	-0.19	fliJ	0.02	yghG	-0.05	pyrB	-0.06	JW5775	0.10
etp	0.29	fliK	0.09	pppA	0.00	pyrL	0.30	JW5776	0.09
yccZ	-0.34	fliL	0.03	glcA	0.03	yjgH	0.03	JW5777	-0.02
ymcA	0.24	fliM	-0.05	glcB	-0.02	yjgI	0.02	JW5778	0.11
ymcB	-0.02	fliN	0.06	glcG	-0.01	yjgJ	0.03	JW5779	0.04
ymcC	-0.02	fliP	-0.08	glcD	-0.05	argI	-0.04	JW5781	0.07
ymcC	-0.14	fliQ	0.07	glcC	0.01	yjgD	-0.03	JW5782	0.06
ymcD	-0.20	fliR	0.06	yghR	-0.03	holC	0.00	JW5783	0.04
cspG	0.32	rcsA	0.02	yghT	-0.04	pepA	0.00	JW5784	0.04
sfa	-0.03	dsrB	-0.10	pitB	0.00	yjgR	0.03	JW5784	-0.03
gnsA	-0.03	yodD	0.08	gsp	-0.02	idnR	0.22	JW5785	-0.01
yccM	0.06	yedP	0.09	hybG	-0.01	idnT	0.37	JW5786	0.08
torS	0.03	JW1939	0.08	hybE	0.01	idnO	0.17	JW5787	0.05
torT	0.04	yodC	0.11	hybC	0.21	idnD	0.24	JW5788	0.10
torR	0.06	yedI	0.01	hybA	-0.01	idnK	0.14	JW5789	0.04
torC	-0.12	yedA	0.06	hyb0	0.02	intB	0.39	JW5790	0.01
torA	-0.15	vsr	0.10	yghW	0.04	yjgW	-0.10	JW5791	0.08
torD	-0.28	dcm	0.10	yghZ	-0.03	yjgX	0.05	JW5792	0.02
yccD	0.05	yedJ	0.03	yqhA	0.08	yjgZ	0.08	JW5793	-0.01
cbpA	0.03	yedR	0.07	yghA	-0.05	yjhE	0.04	JW5794	0.04
yccE	0.04	yedS	-0.01	exbD	0.39	JW4246	-0.02	JW5795	0.04
agp	0.04	hchA	-0.10	exbB	0.20	JW4246	0.01	JW5796	0.08
yccJ	0.04	yedV	0.00	metC	0.00	fecE	0.13	JW5797	0.03
wrbA	0.03	yedW	0.06	yghB	-0.04	fecD	0.12	JW5798	0.01
ymdF	0.03	yedX	-0.07	JW2977	0.16	fecC	-0.04	JW5799	0.04
ycdG	0.07	yedY	-0.05	yqhD	-0.15	fecB	0.16	JW5800	0.18
ycdH	-0.02	yedZ	-0.04	yqhH	0.08	fecA	0.01	JW5801	0.04
ycdH	0.21	yodA	-0.22	sufI	-0.02	fecR	0.04	JW5803	0.12

ycdI	0.05	yodB	0.00	parC	-0.05	fecI	-0.03	JW5805	0.18
ycdJ	0.06	yeeI	0.03	ygiS	0.19	yjhF	0.05	JW5807	-0.02
ycdK	0.07	JW1959	-0.03	ygiU	-0.03	yjhG	0.04	JW5808	0.03
ycdL	-0.03	yeeL	-0.04	ygiV	-0.05	yjhI	0.07	JW5809	0.01
ycdM	-0.03	shiA	-0.14	ygiW	0.08	sgcR	-0.01	JW5811	0.00
ycdC	-0.07	amn	0.05	qseB	0.05	sgcE	0.04	JW5812	0.05
putA	-0.13	yeeN	0.10	qseC	0.03	sgcA	0.02	JW5813	0.06
yzpU	0.02	yeeO	-0.08	ygiZ	0.01	sgcQ	-0.03	JW5814	0.03
ycdN	0.04	cbl	-0.05	mdaB	-0.06	sgcC	0.08	JW5815	-0.03
ycdO	0.07	nac	0.21	ygiN	-0.03	sgcB	0.07	JW5816	-0.06
ycdB	0.01	erfK	0.00	yqiA	-0.06	sgcX	0.03	JW5817	0.02
phoH	-0.18	cobT	0.11	icc	0.10	yjhP	0.04	JW5818	0.03
ycdP	0.00	cobS	0.08	yqiB	0.00	yjhQ	-0.17	JW5819	0.00
ycdQ	-0.02	cobU	-0.15	nudF	0.31	JW4269.5	0.09	JW5820	-0.01
ycdR	-0.01	JW5326	-0.03	tolC	-0.29	yjhR	0.01	JW5821	-0.04
ycdS	-0.02	JW1980	-0.14	ygiC	-0.08	yjhS	0.13	JW5822	0.04
ycdU	-0.05	flu	-0.02	ygiD	-0.09	fimB	-0.03	JW5823	-0.03
ycdW	-0.05	yeeR	-0.12	zupT	-0.02	fimE	-0.05	JW5824	-0.03
ycdX	0.03	yeeS	-0.05	ygiL	-0.01	fimA	0.08	JW5825	0.01
ycdY	-0.02	yeeT	-0.01	yqiG	0.07	fimC	0.03	JW5826	0.08
ycdZ	-0.02	yeeU	0.02	JW3017	0.06	fimD	0.02	JW5827	0.03
csgG	-0.05	yeeV	-0.03	glgS	0.04	fimF	-0.02	JW5828	0.05
csgF	-0.06	yeeW	0.03	yqiJ	-0.17	fimG	-0.07	JW5830	0.02
csgE	0.05	yeeX	0.04	yqiK	-0.04	fimH	0.10	JW5831	0.00
csgD	-0.01	yeeA	0.05	rfaE	-0.08	gntP	0.01	JW5832	0.00
csgB	0.05	gyrI	0.02	glnE	-0.07	uxuA	0.09	JW5833	-0.05
csgA	-0.19	sbcB	0.14	ygiF	-0.07	uxuB	0.26	JW5834	0.02
ymdB	0.00	yeeD	0.09	ygiM	-0.27	uxuR	0.21	JW5836	0.01
ymdB	-0.01	yeeE	0.04	upk	-0.03	yjiC	0.04	JW5837	0.01
mdoC	0.03	yeeY	-0.12	folB	0.06	yjiE	-0.03	JW5838	0.01
mdoG	0.21	yeeZ	0.26	ygiH	-0.03	iadA	0.01	JW5839	0.01
JW1036	0.01	hisL	0.01	ygiP	0.00	yjiG	0.03	JW5840	0.00
mdoH	0.02	hisG	0.06	ttdA	0.36	yjiJ	0.15	JW5841	-0.03
yceK	-0.04	hisD	-0.09	ttdB	0.31	yjiK	0.05	JW5842	-0.01
msyB	-0.10	hisC	-0.07	ygjE	-0.07	JW4296.5	0.04	JW5843	0.04
yceE	0.00	hisB	-0.02	rpsU	-0.04	yjiN	0.05	JW5844	0.00
htrB	0.05	hisA	-0.25	dnaG	-0.14	yjiO	0.01	JW5845	0.00
yceA	0.01	hisF	0.10	rpoD	0.01	yjiP	-0.22	JW5846	0.01
yceA	-0.05	hisI	-0.08	mug	-0.06	yjiR	0.12	JW5847	-0.02
yceI	-0.03	ugd	-0.05	yqjH	-0.02	yjiS	0.08	JW5848	0.00
yceJ	0.00	gnd	0.04	yqjI	-0.05	mcrB	0.05	JW5849	-0.03
yceO	-0.18	wbbL	0.05	aer	-0.03	yjiW	0.04	JW5850	-0.07
solA	0.04	wbbL	-0.05	ygjH	-0.18	hsdS	0.01	JW5851	-0.03
yceP	0.00	yefI	0.25	ebgR	0.17	hsdM	0.75	JW5852	-0.03
dinI	-0.05	wbbJ	-0.02	ebgA	0.08	hsdR	0.06	JW5853	-0.04
pyrC	-0.16	yefG	-0.23	ebgC	0.00	mrr	-0.11	JW5854	-0.02
yceB	-0.01	rfc	0.03	ygjI	0.07	yjiX	0.04	JW5855	0.00
grxB	0.47	glf	0.04	ygjJ	0.11	tsr	-0.07	JW5856	-0.03
yceL	-0.04	rfbX	0.04	ygjK	-0.02	yjjL	0.05	JW5857	-0.06
yceL	-0.03	rfbC	-0.14	fadH	0.06	dnaT	0.03	JW5858	-0.01
JW5859	-0.03	JW5863	-0.02	JW5868	-0.04	JW5872	-0.03	JW5877	-0.03
JW5860	-0.03	JW5864	-0.03	JW5869	-0.03	JW5873	-0.05	JW5878	-0.02
JW5861	0.00	JW5866	-0.05	JW5870	-0.03	JW5874	-0.03	JW5876	-0.05
JW5862	-0.04	JW5867	-0.02	JW5871	-0.04				

10.3 Discussion

We screened the Keio collection of 3985 *E. coli* mutants with in frame deletions of nonessential genes to identify mutants resistant to a T6SS-mediated attack of *V. cholerae* V52. Whereas the focus of our screen was to identify resistant *E. coli* mutants, Tao Dong performed an RNA-Seq analysis on *E. coli* exposed to a T6SS-mediated attack of *V. cholerae* and subsequently identified *E. coli* mutants that are becoming more sensitive to T6SS-mediated killing of *V. cholerae* (Dong et al., 2015). *E. coli* deficient in *soxS* and *soxR*, were killed 3 fold more upon a T6SS-mediated attack of *V. cholerae* than an *E. coli* mutant with a mutation unrelated to the oxidative stress response (Dong et al., 2015). SoxR and SoxS respond to radical oxygen species initiated by contact-dependent and contact-independent perturbations of the outer membrane and induce the redox stress response (Dong et al., 2015). We did not detect a hypersensitivity of *soxR* and *soxS*-deficient mutants in our screen (Table 2). I can imagine two reasons why we did not identify these mutants in our screen. One is that our screen was optimized to detect resistant and not sensitive mutants. Second, differences in killing as small as 3 fold can be detected in the experimental set up of a killing assay in which the CFUs are enumerated after serial dilution of the surviving bacteria, but might not be detected in the experimental setup of a high-throughput screen in which the OD of surviving *E.coli* mutants after exposure to *V. cholerae* was analyzed.

Here, I discuss the implications of the lack of a resistant *E. coli* mutant on the development of therapeutic strategies that target the T6SS. Antibiotics form the basis for many clinical applications such as surgery to prevent and treat bacterial infections (Mertens et al., 1989). An increasing number of bacteria resistant to antibiotics challenge human health (WHO, 2014b). New therapeutic targets and strategies are needed to combat bacterial infections. The T6SS might be a promising new target for a new treatment strategy (S. Pukatzki, unpublished). Compared to targeting house-keeping functions of bacteria such as protein or peptidoglycan synthesis, targeting the T6SS might raise the low numbers of resistance toward the antibacterial agent (S. Pukatzki, unpublished). Antimicrobial agents that target house-keeping functions provide a high selection pressure on bacterial populations. In contrast, the T6SS might only temporarily be exposed to a selection pressure that selects for resistance because the T6SS of

pandemic *V. cholerae* strains is not essential for survival and is activated only under specific conditions (e.g., upon induction by mucin) (S. Pukatzki, unpublished). Experiments in which the occurrence of antimicrobial resistance upon the targeting of house-keeping functions or the T6SS are required to test this concept. The observation that the lack of a single nonessential gene does not render *E. coli* resistant to a T6SS-mediated attack raises the question of whether the T6SS itself could be used as a therapeutic antimicrobial agent. Among the bacteria that are sensitive to T6SS-mediated killing by *V. cholerae* are pathogens such as Salmonella, *enterhaemorrhagic E. coli*, and *enteropathogenic E. coli* (MacIntyre et al., 2010). Further studies are needed to create a T6SS-positive bacterium that is not harmful to the human host but would specifically target a pathogen such as Salmonella and could be used to treat a Salmonella infection. Similar to *E. coli* exposed to the T6SS of *V. cholerae* in our Keio screen, Salmonella would be exposed to multiple T6SS effectors with different antiprokaryotic activities that clear a Salmonella infection without giving Salmonella the chance to become resistant based on modifications of a single, nonessential gene.

11 References

Abdallah, A.M., Gey van Pittius, N.C., Champion, P.A., Cox, J., Luirink, J., Vandenbroucke-Grauls, C.M., Appelmelk, B.J., and Bitter, W. (2007). Type VII secretion--mycobacteria show the way. Nature Reviews Microbiology *5*, 883-891.

Abel, S., Abel zur Wiesch, P., Chang, H.H., Davis, B.M., Lipsitch, M., and Waldor, M.K. (2015). Sequence tag-based analysis of microbial population dynamics. Nature Methods *12*, 223-226, 223 p following 226.

Abuaita, B.H., and Withey, J.H. (2009). Bicarbonate Induces *Vibrio cholerae* virulence gene expression by enhancing ToxT activity. Infection and immunity *77*, 4111-4120.

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. Infection and immunity *78*, 2554-2570.

Almagro-Moreno, S., Pruss, K., and Taylor, R.K. (2015). Intestinal Colonization Dynamics of *Vibrio cholerae*. PLoS Pathogens *11*, e1004787.

Alteri, C.J., Himpsl, S.D., Pickens, S.R., Lindner, J.R., Zora, J.S., Miller, J.E., Arno, P.D., Straight, S.W., and Mobley, H.L. (2013). Multicellular bacteria deploy the type VI secretion system to preemptively strike neighboring cells. PLoS Pathogens *9*, e1003608.

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

Altindis, E., Fu, Y., and Mekalanos, J.J. (2014). Proteomic analysis of *Vibrio cholerae* outer membrane vesicles. Proceedings of the National Academy of Sciences of the United States of America *111*, E1548-1556.

Altschul, S.F., Gish, W., Miller, W., Myers, E.W., and Lipman, D.J. (1990). Basic local alignment search tool. Journal of Molecular Biology *215*, 403-410.

Alvarez-Martinez, C.E., and Christie, P.J. (2009). Biological diversity of prokaryotic type IV secretion systems. Microbiology and Molecular Biology Reviews : MMBR *73*, 775-808.

Aoki, S.K., Diner, E.J., de Roodenbeke, C.T., 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.

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.

Aoki, S.K., Webb, J.S., Braaten, B.A., and Low, D.A. (2009). Contact-dependent growth inhibition causes reversible metabolic downregulation in *Escherichia coli*. Journal of bacteriology *191*, 1777-1786.

Austin, B., and Zhang, X.H. (2006). *Vibrio harveyi*: a significant pathogen of marine vertebrates and invertebrates. Letters in applied microbiology *43*, 119-124.

Baba, T., Ara, T., Hasegawa, M., Takai, Y., Okumura, Y., Baba, M., Datsenko, K.A., Tomita, M., Wanner, B.L., and Mori, H. (2006). Construction of *Escherichia coli* K-12 in-frame, single-gene knockout mutants: the Keio collection. Molecular Systems Biology *2*, 2006 0008.

Babujee, L., Apodaca, J., Balakrishnan, V., Liss, P., Kiley, P.J., Charkowski, A.O., Glasner, J.D., and Perna, N.T. (2012). Evolution of the metabolic and regulatory networks associated with oxygen availability in two phytopathogenic enterobacteria. BMC Genomics *13*, 110.

Barua, D. (1992). History of cholera (Plenum, New York).

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.

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.

Bernard, C.S., Brunet, Y.R., Gavioli, M., Lloubes, R., and Cascales, E. (2011). Regulation of type VI secretion gene clusters by sigma54 and cognate enhancer binding proteins. Journal of Bacteriology *193*, 2158-2167.

Bik, E.M., Bunschoten, A.E., Gouw, R.D., and Mooi, F.R. (1995). Genesis of the novel epidemic *Vibrio cholerae* O139 strain: evidence for horizontal transfer of genes involved in polysaccharide synthesis. The EMBO Journal *14*, 209-216.

Billings J.S., M.E., Peters, J.S. (1975). The cholera epidemic of 1873 in the United States. document 95, 43rd Congress, 2nd session (Washington, D.C.: U.S. Government Printing Office), pp. 1-1025.

Bingle, L.E., Bailey, C.M., and Pallen, M.J. (2008). Type VI secretion: a beginner's guide. Current Opinion in Microbiology *11*, 3-8.

Bishop, A.L., and Camilli, A. (2011). *Vibrio cholerae*: lessons for mucosal vaccine design. Expert review of vaccines *10*, 79-94.

Blattner, F.R., Plunkett, G., 3rd, Bloch, C.A., Perna, N.T., Burland, V., Riley, M., Collado-Vides, J., Glasner, J.D., Rode, C.K., Mayhew, G.F.*, et al.* (1997). The complete genome sequence of *Escherichia coli* K-12. Science *277*, 1453-1462.

Blocker, A., Komoriya, K., and Aizawa, S. (2003). Type III secretion systems and bacterial flagella: insights into their function from structural similarities. Proceedings of the National Academy of Sciences of the United States of America *100*, 3027-3030.

Boardman, B.K., and Satchell, K.J. (2004). *Vibrio cholerae* strains with mutations in an atypical type I secretion system accumulate RTX toxin intracellularly. Journal of bacteriology *186*, 8137-8143.

Bonemann, G., Pietrosiuk, A., Diemand, A., Zentgraf, H., and Mogk, A. (2009). Remodelling of VipA/VipB tubules by ClpV-mediated threading is crucial for type VI protein secretion. The EMBO Journal *28*, 315-325.

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.

Boucher, Y., Cordero, O.X., Takemura, A., Hunt, D.E., Schliep, K., Bapteste, E., Lopez, P., Tarr, C.L., and Polz, M.F. (2011). Local mobile gene pools rapidly cross species boundaries to create endemicity within global *Vibrio cholerae* populations. mBio *2*.

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*, 104.

Bradley, D.E. (1980). Morphological and serological relationships of conjugative pili. Plasmid *4*, 155-169.

Brantl, S. (2012). Bacterial type I toxin-antitoxin systems. RNA biology *9*, 1488-1490.
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. The Journal of Biological Chemistry *288*, 7618-7625.

Brunet, Y.R., Khodr, A., Logger, L., Aussel, L., Mignot, T., Rimsky, S., and Cascales, E. (2015). H-NS silencing of the SPI-6-encoded Type VI secretion system limits *Salmonella enterica* serovar Typhimurium interbacterial killing. Infection and Immunity 83(7):2738-50.

Burrows, W., Mather, A.N., and et al. (1946). Studies on immunity to Asiatic cholera; the O and H antigenic structure of the cholera and related vibrios. The Journal of Infectious diseases *79*, 168-197.

Butler, S.M., and Camilli, A. (2004). Both chemotaxis and net motility greatly influence the infectivity of *Vibrio cholerae*. Proceedings of the National Academy of Sciences of the United States of America *101*, 5018-5023.

Cameron, D.E., Urbach, J.M., and Mekalanos, J.J. (2008). A defined transposon mutant library and its use in identifying motility genes in *Vibrio cholerae*. Proceedings of the National Academy of Sciences of the United States of America *105*, 8736-8741.

Carraro, N., and Burrus, V. (2014). Biology of Three ICE Families: SXT/R391, ICEBs1, and ICESt1/ICESt3. Microbiology Spectrum *2*.

Cascales, E. (2008). The type VI secretion toolkit. EMBO Reports *9*, 735-741.

Cash, R.A., Music, S.I., Libonati, J.P., Craig, J.P., Pierce, N.F., and Hornick, R.B. (1974a). Response of man to infection with *Vibrio cholerae*. II. Protection from illness afforded by previous disease and vaccine. The Journal of Infectious Diseases *130*, 325-333.

Cash, R.A., Music, S.I., Libonati, J.P., Snyder, M.J., Wenzel, R.P., and Hornick, R.B. (1974b). Response of man to infection with *Vibrio cholerae*. I. Clinical, serologic, and bacteriologic responses to a known inoculum. The Journal of Infectious Diseases *129*, 45-52.

Cassel, D., and Pfeuffer, T. (1978). Mechanism of cholera toxin action: covalent modification of the guanyl nucleotide-binding protein of the adenylate cyclase system. Proceedings of the National Academy of Sciences of the United States of America *75*, 2669-2673.

Cassel, D., and Selinger, Z. (1977). Mechanism of adenylate cyclase activation by cholera toxin: inhibition of GTP hydrolysis at the regulatory site. Proceedings of the National Academy of Sciences of the United States of America *74*, 3307-3311.

Cavard, D., Lloubes, R., Morlon, J., Chartier, M., and Lazdunski, C. (1985). Lysis protein encoded by plasmid ColA-CA31. Gene sequence and export. Molecular & general genetics : MGG *199*, 95-100.

Chaand, M., and Dziejman, M. (2013). *Vibrio cholerae* VttR(A) and VttR(B) regulatory influences extend beyond the type 3 secretion system genomic island. Journal of Bacteriology *195*, 2424-2436.

Chaand, M., Miller, K.A., Sofia, M.K., Schlesener, C., Weaver, J.W., Sood, V., and Dziejman, M. (2015). Type 3 Secretion System Island Encoded Proteins Required for Colonization by Non-O1/non-O139 Serogroup *V. cholerae*. Infection and Immunity, IAI.03020-14.

Chambers, J.S. (1938a). Cholera's first invasion of the New World. In the conquest of cholera (New York: MacMillan), pp. 24-44.

Chambers, J.S. (1938b). Pestilence of 1832 along the Atlantic seaboard. The conquest of cholera, 45-84.

Chimetto, L.A., Cleenwerck, I., Alves, N., Jr., Silva, B.S., Brocchi, M., Willems, A., De Vos, P., and Thompson, F.L. (2011). *Vibrio communis* sp. nov., isolated from the marine animals Mussismilia hispida, Phyllogorgia dilatata, Palythoa caribaeorum, Palythoa variabilis and Litopenaeus vannamei. International Journal of Systematic and Evolutionary Microbiology *61*, 362-368.

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*. Proceedings of the National Academy of Sciences of the United States of America *106*, 15442-15447.

Clarholm, M. (1981). Protozoan grazing of bacteria in soil-impact and importance. Microbial Ecology *7*, 343-350.

Clarke, M. (2010). Recent insights into host-pathogen interactions from Dictyostelium. Cellular Microbiology *12*, 283-291.

Clemens, J.D., Sack, D.A., Harris, J.R., Chakraborty, J., Khan, M.R., Huda, S., Ahmed, F., Gomes, J., Rao, M.R., Svennerholm, A.M.*, et al.* (1989). ABO blood groups and cholera: new observations on specificity of risk and modification of vaccine efficacy. The Journal of Infectious Diseases *159*, 770-773.

Clemens, J.D., Sack, D.A., Harris, J.R., Van Loon, F., Chakraborty, J., Ahmed, F., Rao, M.R., Khan, M.R., Yunus, M., Huda, N.*, et al.* (1990). Field trial of oral cholera vaccines in Bangladesh: results from three-year follow-up. Lancet *335*, 270-273.

Cline, K., Ettinger, W.F., and Theg, S.M. (1992). Protein-specific energy requirements for protein transport across or into thylakoid membranes. Two lumenal proteins are transported in the absence of ATP. The Journal of Biological Chemistry *267*, 2688-2696.

Colwell, R.R. (2000). Viable but nonculturable bacteria: a survival strategy. Journal of Infection and Chemotherapy : Official Journal of the Japan Society of Chemotherapy *6*, 121-125.

Colwell, R.R., Brayton, P., Herrington, D., Tall, B., Huq, A., and Levine, M.M. (1996). Viable but non-culturable *Vibrio cholerae* O1 revert to a cultivable state in the human intestine. World Journal of Microbiology & Biotechnology *12*, 28-31.

Costa, T.R., 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. Nature Reviews Microbiology *13*, 343-359.

Daleke, M.H., van der Woude, A.D., Parret, A.H., Ummels, R., de Groot, A.M., Watson, D., Piersma, S.R., Jimenez, C.R., Luirink, J., Bitter, W*., et al.* (2012). Specific chaperones for the type VII protein secretion pathway. The Journal of Biological Chemistry *287*, 31939-31947.

Darling, A.C., Mau, B., Blattner, F.R., and Perna, N.T. (2004). Mauve: multiple alignment of conserved genomic sequence with rearrangements. Genome Research *14*, 1394-1403.

Darriba, D., Taboada, G.L., Doallo, R., and Posada, D. (2012). jModelTest 2: more models, new heuristics and parallel computing. Nature Methods *9*, 772.

De, K., Ramamurthy, T., Faruque, S.M., Yamasaki, S., Takeda, Y., Nair, G.B., and Nandy, R.K. (2004). Molecular characterisation of rough strains of *Vibrio cholerae* isolated from diarrhoeal cases in India and their comparison to smooth strains. FEMS microbiology letters *232*, 23-30.

Desvaux, M., Hebraud, M., Henderson, I.R., and Pallen, M.J. (2006). Type III secretion: what's in a name? Trends in Microbiology *14*, 157-160.

Devault, A.M., Golding, G.B., Waglechner, N., Enk, J.M., Kuch, M., Tien, J.H., Shi, M., Fisman, D.N., Dhody, A.N., Forrest, S*., et al.* (2014). Second-pandemic strain of *Vibrio cholerae* from the Philadelphia cholera outbreak of 1849. The New England Journal of Medicine *370*, 334-340.

Dietz, J., and Field, M. (1973). Ion transport in rabbit ileal mucosa. IV. Bicarbonate secretion. The American Journal of Physiology *225*, 858-861.

Diner, E.J., Beck, C.M., Webb, J.S., Low, D.A., and Hayes, C.S. (2012). Identification of a target cell permissive factor required for contact-dependent growth inhibition (CDI). Genes & Development *26*, 515-525.

DiRita, V.J., and Mekalanos, J.J. (1991). Periplasmic interaction between two membrane regulatory proteins, ToxR and ToxS, results in signal transduction and transcriptional activation. Cell *64*, 29-37.

DiRita, V.J., Parsot, C., Jander, G., and Mekalanos, J.J. (1991). Regulatory cascade controls virulence in *Vibrio cholerae*. Proceedings of the National Academy of Sciences of the United States of America *88*, 5403-5407.

Dong, T.G., Dong, S., Catalano, C., Moore, R., Liang, X., and Mekalanos, J.J. (2015). Generation of reactive oxygen species by lethal attacks from competing microbes.

Proceedings of the National Academy of Sciences of the United States of America *112*, 2181-2186.

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*. Proceedings of the National Academy of Sciences of the United States of America *110*, 2623-2628.

Donia, M.S., Cimermancic, P., Schulze, C.J., Wieland Brown, L.C., Martin, J., Mitreva, M., Clardy, J., Linington, R.G., and Fischbach, M.A. (2014). A systematic analysis of biosynthetic gene clusters in the human microbiome reveals a common family of antibiotics. Cell *158*, 1402-1414.

Duan, F., and March, J.C. (2010). Engineered bacterial communication prevents *Vibrio cholerae* virulence in an infant mouse model. Proceedings of the National Academy of Sciences of the United States of America *107*, 11260-11264.

Durand, E., Derrez, E., Audoly, G., Spinelli, S., Ortiz-Lombardia, M., Raoult, D., Cascales, E., and Cambillau, C. (2012a). Crystal structure of the VgrG1 actin cross-linking domain of the *Vibrio cholerae* type VI secretion system. The Journal of Biological Chemistry *287*, 38190-38199.

Durand, E., Zoued, A., Spinelli, S., Watson, P.J., Aschtgen, M.S., Journet, L., Cambillau, C., and Cascales, E. (2012b). Structural characterization and oligomerization of the TssL protein, a component shared by bacterial type VI and type IVb secretion systems. The Journal of Biological Chemistry *287*, 14157-14168.

Durrenberger, M.B., Villiger, W., and Bachi, T. (1991). Conjugational junctions: morphology of specific contacts in conjugating *Escherichia coli* bacteria. Journal of Structural Biology *107*, 146-156.

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. Proceedings of the National Academy of Sciences of the United States of America *102*, 3465-3470.

Edgar, R.C. (2004). MUSCLE: multiple sequence alignment with high accuracy and high throughput. Nucleic Acids Research *32*, 1792-1797.

Elliott, S.J., and Kaper, J.B. (1997). Role of type 1 fimbriae in EPEC infections. Microbial pathogenesis *23*, 113-118.

Faruque, S.M., Albert, M.J., and Mekalanos, J.J. (1998). Epidemiology, genetics, and ecology of toxigenic *Vibrio cholerae*. Microbiology and Molecular Biology Reviews : MMBR *62*, 1301-1314.

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. Proceedings of the National Academy of Sciences of the United States of America *101*, 2123-2128.

Faruque, S.M., Islam, M.J., Ahmad, Q.S., Biswas, K., Faruque, A.S., Nair, G.B., Sack, R.B., Sack, D.A., and Mekalanos, J.J. (2006). An improved technique for isolation of environmental *Vibrio cholerae* with epidemic potential: monitoring the emergence of a multiple-antibiotic-resistant epidemic strain in Bangladesh. The Journal of Infectious Diseases *193*, 1029-1036.

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

Faruque, S.M., Naser, I.B., Islam, M.J., Faruque, A.S., Ghosh, A.N., Nair, G.B., Sack, D.A., and Mekalanos, J.J. (2005). Seasonal epidemics of cholera inversely correlate with the prevalence of environmental cholera phages. Proceedings of the National Academy of Sciences of the United States of America *102*, 1702-1707.

Faruque, S.M., Sack, D.A., Sack, R.B., Colwell, R.R., Takeda, Y., and Nair, G.B. (2003b). Emergence and evolution of *Vibrio cholerae* O139. Proceedings of the National Academy of Sciences of the United States of America *100*, 1304-1309.

Faruque, S.M., Zhu, J., Asadulghani, Kamruzzaman, M., and Mekalanos, J.J. (2003c). Examination of diverse toxin-coregulated pilus-positive *Vibrio cholerae* strains fails to demonstrate evidence for *Vibrio* pathogenicity island phage. Infection and immunity *71*, 2993-2999.

Felisberto-Rodrigues, C., Durand, E., Aschtgen, M.S., Blangy, S., Ortiz-Lombardia, M., Douzi, B., Cambillau, C., and Cascales, E. (2011). Towards a structural comprehension of bacterial type VI secretion systems: characterization of the TssJ-TssM complex of an *Escherichia coli* pathovar. PLoS Pathogens *7*, e1002386.

Feulner, G., Gray, J.A., Kirschman, J.A., Lehner, A.F., Sadosky, A.B., Vlazny, D.A., Zhang, J., Zhao, S., and Hill, C.W. (1990). Structure of the rhsA locus from *Escherichia coli* K-12 and comparison of rhsA with other members of the rhs multigene family. Journal of Bacteriology *172*, 446-456.

Finkelstein, R.A. (1962). Vibriocidal antibody inhibition (VAI) analysis: a technique for the identification of the predominant vibriocidal antibodies in serum and for the detection and identification of *Vibrio cholerae* antigens. J Immunol *89*, 264-271.

Flach, C.F., Qadri, F., Bhuiyan, T.R., Alam, N.H., Jennische, E., Lonnroth, I., and Holmgren, J. (2007). Broad up-regulation of innate defense factors during acute cholera. Infection and Immunity *75*, 2343-2350.

Frobel, J., Rose, P., Lausberg, F., Blummel, A.S., Freudl, R., and Muller, M. (2012). Transmembrane insertion of twin-arginine signal peptides is driven by TatC and regulated by TatB. Nature Communications *3*, 1311.

Fu, Y., Waldor, M.K., and Mekalanos, J.J. (2013). Tn-Seq analysis of *Vibrio cholerae* intestinal colonization reveals a role for T6SS-mediated antibacterial activity in the host. Cell Host & Microbe *14*, 652-663.

Gangarosa, E.J.a.W.H.M. (1974). Epidemiology and surveillance of cholera. In Cholera, D.B.a.W. Burrows, ed. (Philadelphia: W. B. Saunders Co.).

Gardel, C.L., and Mekalanos, J.J. (1996). Alterations in *Vibrio cholerae* motility phenotypes correlate with changes in virulence factor expression. Infection and Immunity *64*, 2246-2255.

Gibbs, K.A., Urbanowski, M.L., and Greenberg, E.P. (2008). Genetic determinants of self identity and social recognition in bacteria. Science *321*, 256-259.

Gill, D.M., and King, C.A. (1975). The mechanism of action of cholera toxin in pigeon erythrocyte lysates. The Journal of Biological Chemistry *250*, 6424-6432.

Gill, D.M., and Meren, R. (1978). ADP-ribosylation of membrane proteins catalyzed by cholera toxin: basis of the activation of adenylate cyclase. Proceedings of the National Academy of Sciences of the United States of America *75*, 3050-3054.

Gonzales, M.F., Brooks, T., Pukatzki, S.U., and Provenzano, D. (2013). Rapid protocol for preparation of electrocompetent *Escherichia coli* and *Vibrio cholerae*. Journal of Visualized Experiments : JoVE 8(80).

Govan, J.R., and Deretic, V. (1996). Microbial pathogenesis in cystic fibrosis: mucoid *Pseudomonas aeruginosa* and *Burkholderia cepacia*. Microbiological Reviews *60*, 539-574.

Gray, M.D., Bagdasarian, M., Hol, W.G., and Sandkvist, M. (2011). In vivo cross-linking of EpsG to EpsL suggests a role for EpsL as an ATPase-pseudopilin coupling protein in the Type II secretion system of *Vibrio cholerae*. Molecular Microbiology *79*, 786-798.

Groth, A.C., and Calos, M.P. (2004). Phage integrases: biology and applications. Journal of Molecular Biology *335*, 667-678.

Guentzel, M.N., and Berry, L.J. (1975). Motility as a virulence factor for *Vibrio cholerae*. Infection and Immunity *11*, 890-897.

Guindon, S., and Gascuel, O. (2003). A simple, fast, and accurate algorithm to estimate large phylogenies by maximum likelihood. Systematic Biology *52*, 696-704.

Guinn, K.M., Hickey, M.J., Mathur, S.K., Zakel, K.L., Grotzke, J.E., Lewinsohn, D.M., Smith, S., and Sherman, D.R. (2004). Individual RD1-region genes are required for export of ESAT-6/CFP-10 and for virulence of *Mycobacterium tuberculosis*. Molecular Microbiology *51*, 359-370.

Gupta, S., and Chowdhury, R. (1997). Bile affects production of virulence factors and motility of *Vibrio cholerae*. Infection and Immunity *65*, 1131-1134.

Guzman, L.M., Belin, D., Carson, M.J., and Beckwith, J. (1995). Tight regulation, modulation, and high-level expression by vectors containing the arabinose PBAD promoter. Journal of Bacteriology *177*, 4121-4130.

Hall, R.M., and Collis, C.M. (1995). Mobile gene cassettes and integrons: capture and spread of genes by site-specific recombination. Molecular Microbiology *15*, 593-600.

Harris, A.M., Bhuiyan, M.S., Chowdhury, F., Khan, A.I., Hossain, A., Kendall, E.A., Rahman, A., LaRocque, R.C., Wrammert, J., Ryan, E.T.*, et al.* (2009). Antigen-specific memory B-cell responses to *Vibrio cholerae* O1 infection in Bangladesh. Infection and Immunity *77*, 3850-3856.

Harris, J.B., Khan, A.I., LaRocque, R.C., Dorer, D.J., Chowdhury, F., Faruque, A.S., Sack, D.A., Ryan, E.T., Qadri, F., and Calderwood, S.B. (2005). Blood group, immunity, and risk of infection with *Vibrio cholerae* in an area of endemicity. Infection and Immunity *73*, 7422-7427.

Harrison, J.W., Dung, T.T., Siddiqui, F., Korbrisate, S., Bukhari, H., Tra, M.P., Hoang, N.V., Carrique-Mas, J., Bryant, J., Campbell, J.I.*, et al.* (2014). Identification of possible virulence marker from *Campylobacter jejuni* isolates. Emerging Infectious Diseases *20*, 1026-1029.

Hase, C.C., and Finkelstein, R.A. (1991). Cloning and nucleotide sequence of the *Vibrio cholerae* hemagglutinin/protease (HA/protease) gene and construction of an HA/protease-negative strain. Journal of Bacteriology *173*, 3311-3317.

Hays, J.N. (2005). Epidemics and Pandemics: Their Impacts on Human History (ABC-CLIO).

Heidelberg, J.F., Eisen, 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.

Herrington, D.A., Hall, R.H., Losonsky, G., Mekalanos, J.J., Taylor, R.K., and Levine, M.M. (1988). Toxin, toxin-coregulated pili, and the toxR regulon are essential for *Vibrio cholerae* pathogenesis in humans. The Journal of experimental medicine *168*, 1487-1492.

Hibbing, M.E., Fuqua, C., Parsek, M.R., and Peterson, S.B. (2010). Bacterial competition: surviving and thriving in the microbial jungle. Nature Reviews Microbiology *8*, 15-25.

Higgins, D.E., Nazareno, E., and DiRita, V.J. (1992). The virulence gene activator ToxT from *Vibrio cholerae* is a member of the AraC family of transcriptional activators. Journal of Bacteriology *174*, 6974-6980.

Hill, C.W. (1999). Large genomic sequence repetitions in bacteria: lessons from rRNA operons and Rhs elements. Research in Microbiology *150*, 665-674.

Hill, C.W., Feulner, G., Brody, M.S., Zhao, S., Sadosky, A.B., and Sandt, C.H. (1995). Correlation of Rhs elements with *Escherichia coli* population structure. Genetics *141*, 15-24.

Ho, B.T., Basler, M., and Mekalanos, J.J. (2013). Type 6 secretion system-mediated immunity to type 4 secretion system-mediated gene transfer. Science *342*, 250-253.

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.

Hollenstein, K., Dawson, R.J., and Locher, K.P. (2007). Structure and mechanism of ABC transporter proteins. Current Opinion in Structural Biology *17*, 412-418.

Holmgren, J. (1981). Actions of cholera toxin and the prevention and treatment of cholera. Nature *292*, 413-417.

Horvath, P., and Barrangou, R. (2010). CRISPR/Cas, the immune system of bacteria and archaea. Science *327*, 167-170.

Hsiao, A., Ahmed, A.M., Subramanian, S., Griffin, N.W., Drewry, L.L., Petri, W.A., Jr., Haque, R., Ahmed, T., and Gordon, J.I. (2014). Members of the human gut microbiota involved in recovery from *Vibrio cholerae* infection. Nature *515*, 423-426.

Huq, A., Xu, B., Chowdhury, M.A., Islam, M.S., Montilla, R., and Colwell, R.R. (1996). A simple filtration method to remove plankton-associated *Vibrio cholerae* in raw water supplies in developing countries. Applied and Environmental Microbiology *62*, 2508-2512.

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.* Infection and Immunity *55*, 1090-1093.

Iredell, J.R., and Manning, P.A. (1994). The toxin-co-regulated pilus of *Vibrio cholerae* O1: a model for type 4 pilus biogenesis? Trends in Microbiology *2*, 187-192.

Ishikawa, T., Sabharwal, D., Broms, J., Milton, D.L., Sjostedt, A., Uhlin, B.E., and Wai, S.N. (2012). Pathoadaptive conditional regulation of the type VI secretion system in *Vibrio cholerae* O1 strains. Infection and Immunity *80*, 575-584.

Isnansetyo, A., and Kamei, Y. (2003). *Pseudoalteromonas phenolica* sp. nov., a novel marine bacterium that produces phenolic anti-methicillin-resistant *Staphylococcus aureus* substances. International Journal of Systematic and Evolutionary Microbiology *53*, 583-588.

Israil, A.M., Nacescu, N., Ciufecu, C., and Cedru, C. (1987). The development and application of a bacteriocinogenotyping scheme for *Vibrio cholerae* non-group O-1 strains. Zentralblatt fur Bakteriologie, Mikrobiologie, und Hygiene Series A, Medical microbiology, infectious diseases, virology, parasitology *264*, 235-245.

Israil, A.M., Nacescu, N., Ciufecu, C., and Stefanescu, C. (1983). Studies on bacteriocin production by NAG-strains of *Vibrio cholerae* as a possible epidemiological marker. Zentralblatt fur Bakteriologie, Mikrobiologie und Hygiene 1 Abt Originale A, Medizinische Mikrobiologie, Infektionskrankheiten und Parasitologie = International Journal of Microbiology and Hygiene A, Medical micro *255*, 285-293.

Iwanaga, M., Yamamoto, K., Higa, N., Ichinose, Y., Nakasone, N., and Tanabe, M. (1986). Culture conditions for stimulating cholera toxin production by *Vibrio cholerae* O1 El Tor. Microbiology and Immunology *30*, 1075-1083.

Jardetzky, O. (1966). Simple allosteric model for membrane pumps. Nature *211*, 969-970.

Jiang, F., Waterfield, N.R., Yang, J., Yang, G., and Jin, Q. (2014). A *Pseudomonas aeruginosa* type VI secretion phospholipase D effector targets both prokaryotic and eukaryotic cells. Cell Host & Microbe *15*, 600-610.

Jones, D., Metzger, H.J., Schatz, A., and Waksman, S.A. (1944). Control of Gram-Negative Bacteria in Experimental Animals by Streptomycin. Science *100*, 103-105.

Jones, P.M., O'Mara, M.L., and George, A.M. (2009). ABC transporters: a riddle wrapped in a mystery inside an enigma. Trends in Biochemical Sciences *34*, 520-531.

Jonson, G., A.-M. Svennerholm, and J. Holmgren (1990). Expression of virulence factors by classical and El Tor *Vibrio cholerae* in vivo and in vitro. FEMS Microbiol Ecol *74*, 221-228.

Judson, N., and Mekalanos, J.J. (2000). TnAraOut, a transposon-based approach to identify and characterize essential bacterial genes. Nature Biotechnology *18*, 740-745.

Jurenaite, M., Markuckas, A., and Suziedeliene, E. (2013). Identification and characterization of type II toxin-antitoxin systems in the opportunistic pathogen *Acinetobacter baumannii*. Journal of Bacteriology *195*, 3165-3172.

Kahn, R.A., and Gilman, A.G. (1984). ADP-ribosylation of Gs promotes the dissociation of its alpha and beta subunits. The Journal of Biological Chemistry *259*, 6235-6240.

Kamada, N., Kim, Y.G., Sham, H.P., Vallance, B.A., Puente, J.L., Martens, E.C., and Nunez, G. (2012). Regulated virulence controls the ability of a pathogen to compete with the gut microbiota. Science *336*, 1325-1329.

Kamal, A.M. (1974). The seventh pandemic of cholera. (Philadelphia: W. B. Saunders).
Kamp, H.D., Patimalla-Dipali, B., Lazinski, D.W., Wallace-Gadsden, F., and Camilli, A. (2013). Gene fitness landscapes of *Vibrio cholerae* at important stages of its life cycle. PLoS Pathogens *9*, e1003800.

Kanonenberg, K., Schwarz, C.K., and Schmitt, L. (2013). Type I secretion systems - a story of appendices. Research in Microbiology *164*, 596-604.

Kaper, J.B., Morris, J.G., Jr., and Levine, M.M. (1995). Cholera. Clinical Microbiology reviews *8*, 48-86.

Karamanou, S., Vrontou, E., Sianidis, G., Baud, C., Roos, T., Kuhn, A., Politou, A.S., and Economou, A. (1999). A molecular switch in SecA protein couples ATP hydrolysis to protein translocation. Molecular Microbiology *34*, 1133-1145.

Karaolis, D.K., Somara, S., Maneval, D.R., Jr., Johnson, J.A., and Kaper, J.B. (1999). A bacteriophage encoding a pathogenicity island, a type-IV pilus and a phage receptor in cholera bacteria. Nature *399*, 375-379.

Kearse, M., Moir, R., Wilson, A., Stones-Havas, S., Cheung, M., Sturrock, S., Buxton, S., Cooper, A., Markowitz, S., Duran, C.*, et al.* (2012). Geneious Basic: an integrated and extendable desktop software platform for the organization and analysis of sequence data. Bioinformatics *28*, 1647-1649.

Kell, D.B., Kaprelyants, A.S., Weichart, D.H., Harwood, C.R., and Barer, M.R. (1998). Viability and activity in readily culturable bacteria: a review and discussion of the practical issues. Antonie van Leeuwenhoek *73*, 169-187.

Kelley, L.A., Mezulis, S., Yates, C.M., Wass, M.N., and Sternberg, M.J. (2015). The Phyre2 web portal for protein modeling, prediction and analysis. Nature Protocols *10*, 845-858.

Kim, H.B., Wang, M., Ahmed, S., Park, C.H., LaRocque, R.C., Faruque, A.S., Salam, M.A., Khan, W.A., Qadri, F., Calderwood, S.B.*, et al.* (2010). Transferable quinolone resistance in *Vibrio cholerae*. Antimicrobial Agents and Chemotherapy *54*, 799-803.

King, C.A., and Van Heyningen, W.E. (1973). Deactivation of cholera toxin by a sialidase-resistant monosialosylganglioside. The Journal of Infectious Diseases *127*, 639-647.

Kirchberger, P.C., Turnsek, M., Hunt, D.E., Haley, B.J., Colwell, R.R., Polz, M.F., Tarr, C.L., and Boucher, Y. (2014). *Vibrio metoecus* sp. nov., a close relative of *Vibrio cholerae* isolated from coastal brackish ponds and clinical specimens. International journal of systematic and evolutionary microbiology *64*, 3208-3214.

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

Kitaoka, M., Miyata, S.T., Brooks, T.M., Unterweger, D., and Pukatzki, S. (2011a). VasH is a transcriptional regulator of the type VI secretion system functional in endemic and pandemic *Vibrio cholerae*. Journal of Bacteriology *193*, 6471-6482.

Kitaoka, M., Miyata, S.T., Unterweger, D., and Pukatzki, S. (2011b). Antibiotic resistance mechanisms of *Vibrio cholerae*. Journal of Medical Microbiology *60*, 397-407.
Koch, R. (1894). An address on cholera and its bacillus. Br Med J *2*, 453-459.

Korotkov, K.V., Sandkvist, M., and Hol, W.G. (2012). The type II secretion system: biogenesis, molecular architecture and mechanism. Nature Reviews Microbiology *10*, 336-351.

Koskiniemi, S., Garza-Sanchez, F., Sandegren, L., Webb, J.S., Braaten, B.A., Poole, S.J., Andersson, D.I., Hayes, C.S., and Low, D.A. (2014). Selection of orphan Rhs toxin expression in evolved *Salmonella enterica* serovar Typhimurium. PLoS Genetics *10*, e1004255.

Koskiniemi, S., Lamoureux, J.G., Nikolakakis, K.C., t'Kint de Roodenbeke, C., Kaplan, M.D., Low, D.A., and Hayes, C.S. (2013). Rhs proteins from diverse bacteria mediate intercellular competition. Proceedings of the National Academy of Sciences of the United States of America *110*, 7032-7037.

Krebs, S.J., and Taylor, R.K. (2011). Protection and attachment of *Vibrio cholerae* mediated by the toxin-coregulated pilus in the infant mouse model. Journal of bacteriology *193*, 5260-5270.

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 Reports *8*, 20-30.

Kudryashev, M., Wang, R.Y., Brackmann, M., Scherer, S., Maier, T., Baker, D., DiMaio, F., Stahlberg, H., Egelman, E.H., and Basler, M. (2015). Structure of the type VI secretion system contractile sheath. Cell *160*, 952-962.

Lara-Tejero, M., Kato, J., Wagner, S., Liu, X., and Galan, J.E. (2011). A sorting platform determines the order of protein secretion in bacterial type III systems. Science *331*, 1188-1191.

Larkin, M.A., Blackshields, G., Brown, N.P., Chenna, R., McGettigan, P.A., McWilliam, H., Valentin, F., Wallace, I.M., Wilm, A., Lopez, R.*, et al.* (2007). Clustal W and Clustal X version 2.0. Bioinformatics *23*, 2947-2948.

Laval, E. (1989). El colera en Chile (1886-1888). Rev Chil Infect *6*, 96-99.

Lazdunski, C.J., Bouveret, E., Rigal, A., Journet, L., Lloubes, R., and Benedetti, H. (1998). Colicin import into Escherichia coli cells. Journal of Bacteriology *180*, 4993-5002.

Lee, S.H., Butler, S.M., and Camilli, A. (2001). Selection for in vivo regulators of bacterial virulence. Proceedings of the National Academy of Sciences of the United States of America *98*, 6889-6894.

Lee, S.H., Hava, D.L., Waldor, M.K., and Camilli, A. (1999). Regulation and temporal expression patterns of *Vibrio cholerae* virulence genes during infection. Cell *99*, 625-634.
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. Proceedings of the National Academy of Sciences of the United States of America *106*, 4154-4159.

Leiman, P.G., and Shneider, M.M. (2012). Contractile tail machines of bacteriophages. Advances in Experimental Medicine and Biology *726*, 93-114.

Leo, J.C., Grin, I., and Linke, D. (2012). Type V secretion: mechanism(s) of autotransport through the bacterial outer membrane. Philosophical transactions of the Royal Society of London Series B, Biological Sciences *367*, 1088-1101.

LeRoux, M., Kirkpatrick, R.L., Montauti, E.I., Tran, B.Q., Peterson, S.B., Harding, B.N., Whitney, J.C., Russell, A.B., Traxler, B., Goo, Y.A.*, et al.* (2015). Kin cell lysis is a danger signal that activates antibacterial pathways of *Pseudomonas aeruginosa*. eLife *4*.

Levine, M.M., Black, R.E., Clements, M.L., Cisneros, L., Nalin, D.R., and Young, C.R. (1981). Duration of infection-derived immunity to cholera. The Journal of Infectious diseases *143*, 818-820.

Levine, M.M., Kaper, J.B., Herrington, D., Ketley, J., Losonsky, G., Tacket, C.O., Tall, B., and Cryz, S. (1988a). Safety, immunogenicity, and efficacy of recombinant live oral cholera vaccines, CVD 103 and CVD 103-HgR. Lancet *2*, 467-470.

Levine, M.M., Kaper, J.B., Herrington, D., Losonsky, G., Morris, J.G., Clements, M.L., Black, R.E., Tall, B., and Hall, R. (1988b). Volunteer studies of deletion mutants of *Vibrio cholerae* O1 prepared by recombinant techniques. Infection and Immunity *56*, 161-167.

Levine, M.M., Nalin, D.R., Craig, J.P., Hoover, D., Bergquist, E.J., Waterman, D., Holley, H.P., Hornick, R.B., Pierce, N.P., and Libonati, J.P. (1979). Immunity of cholera in man: relative role of antibacterial versus antitoxic immunity. Transactions of the Royal Society of Tropical Medicine and Hygiene *73*, 3-9.

Li, M., Le Trong, I., Carl, M.A., Larson, E.T., Chou, S., De Leon, J.A., Dove, S.L., Stenkamp, R.E., and Mougous, J.D. (2012). Structural basis for type VI secretion effector recognition by a cognate immunity protein. PLoS Pathogens *8*, e1002613.

Lin, R.J., Capage, M., and Hill, C.W. (1984). A repetitive DNA sequence, rhs, responsible for duplications within the Escherichia coli K-12 chromosome. Journal of Molecular Biology *177*, 1-18.

Lipp, E.K., Huq, A., and Colwell, R.R. (2002). Effects of global climate on infectious disease: the cholera model. Clinical Microbiology Reviews *15*, 757-770.

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

Lombardo, M.J., Michalski, J., Martinez-Wilson, H., Morin, C., Hilton, T., Osorio, C.G., Nataro, J.P., Tacket, C.O., Camilli, A., and Kaper, J.B. (2007). An in vivo expression technology screen for *Vibrio cholerae* genes expressed in human volunteers. Proceedings of the National Academy of Sciences of the United States of America *104*, 18229-18234.

Lycklama, A.N.J.A., and Driessen, A.J. (2012). The bacterial Sec-translocase: structure and mechanism. Philosophical transactions of the Royal Society of London Series B, Biological Sciences *367*, 1016-1028.

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

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. Proceedings of the National Academy of Sciences of the United States of America *107*, 4365-4370.

Ma, L.S., Hachani, A., Lin, J.S., Filloux, A., and Lai, E.M. (2014a). *Agrobacterium tumefaciens* Deploys a Superfamily of Type VI Secretion DNase Effectors as Weapons for Interbacterial Competition In Planta. Cell Host & Microbe.

Ma, L.S., Hachani, A., Lin, J.S., Filloux, A., and Lai, E.M. (2014b). Agrobacterium tumefaciens deploys a superfamily of type VI secretion DNase effectors as weapons for interbacterial competition in planta. Cell Host & Microbe *16*, 94-104.

Ma, L.S., Lin, J.S., and Lai, E.M. (2009b). An IcmF family protein, ImpLM, is an integral inner membrane protein interacting with ImpKL, and its walker a motif is required for type VI secretion system-mediated Hcp secretion in *Agrobacterium tumefaciens*. Journal of Bacteriology *191*, 4316-4329.

Ma, L.S., Narberhaus, F., and Lai, E.M. (2012). IcmF family protein TssM exhibits ATPase activity and energizes type VI secretion. The Journal of Biological Chemistry *287*, 15610-15621.

MacIntyre, D.L., Miyata, S.T., Kitaoka, M., and Pukatzki, S. (2010). The *Vibrio cholerae* type VI secretion system displays antimicrobial properties. Proceedings of the National Academy of Sciences of the United States of America *107*, 19520-19524.

Madigan, M., and Martinko, J. (2006). Brock biology of microorganisms, Vol 1 (Pearson Prentice Hall).

Mandlik, A., Livny, J., Robins, W.P., Ritchie, J.M., Mekalanos, J.J., and Waldor, M.K. (2011). RNA-Seq-based monitoring of infection-linked changes in *Vibrio cholerae* gene expression. Cell Host & Microbe *10*, 165-174.

Maniatis (1987). Molecular Cloning, A Laboratory manual, 1 edn.

Mansfield, J., Genin, S., Magori, S., Citovsky, V., Sriariyanum, M., Ronald, P., Dow, M., Verdier, V., Beer, S.V., Machado, M.A.*, et al.* (2012). Top 10 plant pathogenic bacteria in molecular plant pathology. Molecular Plant Pathology *13*, 614-629.

Mazel, D., Dychinco, B., Webb, V.A., and Davies, J. (1998). A distinctive class of integron in the *Vibrio cholerae* genome. Science *280*, 605-608.

McCormack, W.M., Chakraborty, J., Rahman, A.S., and Mosley, W.H. (1969). Vibriocidal antibody in clinical cholera. The Journal of Infectious Diseases *120*, 192-201.

Megli, C.J., Yuen, A.S., Kolappan, S., Richardson, M.R., Dharmasena, M.N., Krebs, S.J., Taylor, R.K., and Craig, L. (2011). Crystal structure of the *Vibrio cholerae* colonization factor TcpF and identification of a functional immunogenic site. Journal of Molecular Biology *409*, 146-158.

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.

Mekalanos, J. (1995). Live cholera vaccines: perspectives on their construction and safety. Bull Inst Pasteur *93*, 255-262.

Mekalanos, J.J. (1983). Duplication and amplification of toxin genes in *Vibrio cholerae*. Cell *35*, 253-263.

Merrell, D.S., Butler, S.M., Qadri, F., Dolganov, N.A., Alam, A., Cohen, M.B., Calderwood, S.B., Schoolnik, G.K., and Camilli, A. (2002a). Host-induced epidemic spread of the cholera bacterium. Nature *417*, 642-645.

Merrell, D.S., and Camilli, A. (1999). The cadA gene of *Vibrio cholerae* is induced during infection and plays a role in acid tolerance. Molecular microbiology *34*, 836-849.
Merrell, D.S., Hava, D.L., and Camilli, A. (2002b). Identification of novel factors involved in colonization and acid tolerance of *Vibrio cholerae*. Molecular Microbiology *43*, 1471-1491.

Mertens, R., Verbist, L., Gordts, B., Lauwers, S., Potvliege, C., Reybrouck, G., Verschraegen, G., Wauters, G., Berghmans, L., Dondeyne, F*., et al.* (1989). National study on the utilization of prophylactic antibiotics in surgery, Belgium, 1986. Epidemiology and Infection *103*, 311-322.

Metcalf, W.W., Jiang, W., Daniels, L.L., Kim, S.K., Haldimann, A., and Wanner, B.L. (1996). Conditionally replicative and conjugative plasmids carrying lacZ alpha for cloning, mutagenesis, and allele replacement in bacteria. Plasmid *35*, 1-13.

Metzger, L.C., and Blokesch, M. (2014). Composition of the DNA-uptake complex of. Mobile Genetic Elements *4*, e28142.

Miller, C.E., Wong, K.H., Feeley, J.C., and Forlines, M.E. (1972). Immunological conversion of *Vibrio cholerae* in gnotobiotic mice. Infection and Immunity *6*, 739-742.

Miller, V.L., and Mekalanos, J.J. (1984). Synthesis of cholera toxin is positively regulated at the transcriptional level by toxR. Proceedings of the National Academy of Sciences of the United States of America *81*, 3471-3475.

Miller, V.L., Taylor, R.K., and Mekalanos, J.J. (1987). Cholera toxin transcriptional activator toxR is a transmembrane DNA binding protein. Cell *48*, 271-279.

Millet, Y.A., Alvarez, D., Ringgaard, S., von Andrian, U.H., Davis, B.M., and Waldor, M.K. (2014). Insights into *Vibrio cholerae* intestinal colonization from monitoring fluorescently labeled bacteria. PLoS pathogens *10*, e1004405.

Mitra, R.K., Nandy, R.K., Ramamurthy, T., Bhattacharya, S.K., Yamasaki, S., Shimada, T., Takeda, Y., and Nair, G.B. (2001). Molecular characterisation of rough variants of *Vibrio cholerae* isolated from hospitalised patients with diarrhoea. Journal of Medical Microbiology *50*, 268-276.

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*. Infection and Immunity *79*, 2941-2949.

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 Pathogens *9*, e1003752.

Moseley, S.L., and Falkow, S. (1980). Nucleotide sequence homology between the heat-labile enterotoxin gene of *Escherichia coli* and *Vibrio cholerae* deoxyribonucleic acid. Journal of Bacteriology *144*, 444-446.

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

Mukerjee, S. (1963). Bacteriophage typing of cholera. Bulletin of the World Health Organization *28*, 337-345.

Muller, M. (2005). Twin-arginine-specific protein export in *Escherichia coli*. Research in Microbiology *156*, 131-136.

Murdoch, S.L., Trunk, K., English, G., Fritsch, M.J., Pourkarimi, E., and Coulthurst, S.J. (2011). The opportunistic pathogen Serratia marcescens utilizes type VI secretion to target bacterial competitors. Journal of Bacteriology *193*, 6057-6069.

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.

Nair, G.B., Faruque, S.M., Bhuiyan, N.A., Kamruzzaman, M., Siddique, A.K., and Sack, D.A. (2002). New variants of *Vibrio cholerae* O1 biotype El Tor with attributes of the classical biotype from hospitalized patients with acute diarrhea in Bangladesh. Journal of Clinical Microbiology *40*, 3296-3299.

Nelson, E.J., Chowdhury, A., Flynn, J., Schild, S., Bourassa, L., Shao, Y., LaRocque, R.C., Calderwood, S.B., Qadri, F., and Camilli, A. (2008). Transmission of *Vibrio cholerae* is antagonized by lytic phage and entry into the aquatic environment. PLoS pathogens *4*, e1000187.

Neoh, S.H., and Rowley, D. (1970). The antigens of *Vibrio cholerae* involved in the vibriocidal action of antibody and complement. The Journal of Infectious Diseases *121*, 505-513.

Newsom, S.W. (2006). Pioneers in infection control: John Snow, Henry Whitehead, the Broad Street pump, and the beginnings of geographical epidemiology. The Journal of Hospital Infection *64*, 210-216.

Ng, W.L., and Bassler, B.L. (2009). Bacterial quorum-sensing network architectures. Annual Review of Genetics *43*, 197-222.

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 Pathogens *6*, e1001102.

Nivaskumar, M., and Francetic, O. (2014). Type II secretion system: a magic beanstalk or a protein escalator. Biochimica et Biophysica Acta *1843*, 1568-1577.

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.

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

Orata, F.D., Keim, P.S., and Boucher, Y. (2014). The 2010 cholera outbreak in Haiti: how science solved a controversy. PLoS Pathogens *10*, e1003967.

Pacini, F. (1854). Osservazioni microscopiche e deduzioni patologiche sul cholera asiatico. Gaz Med Ital Toscana *6*, 397-405.

Paetzel, M., Karla, A., Strynadka, N.C., and Dalbey, R.E. (2002). Signal peptidases. Chemical Reviews *102*, 4549-4580.

Page, A.L., and Parsot, C. (2002). Chaperones of the type III secretion pathway: jacks of all trades. Molecular Microbiology *46*, 1-11.

PAHO (2001). Zoonoses and communicable diseases common to man and animals. In, P.a.H. Organization, ed. (Washington, D.C.).

Palmer, T., and Berks, B.C. (2012). The twin-arginine translocation (Tat) protein export pathway. Nature Reviews Microbiology *10*, 483-496.

Papanikou, E., Karamanou, S., and Economou, A. (2007). Bacterial protein secretion through the translocase nanomachine. Nature Reviews Microbiology *5*, 839-851.

Patel, R., Smith, S.M., and Robinson, C. (2014). Protein transport by the bacterial Tat pathway. Biochimica et Biophysica Acta *1843*, 1620-1628.

Pearson, W.R. (2013). An introduction to sequence similarity ("homology") searching. Current protocols in bioinformatics / editoral board, Andreas D Baxevanis [et al] *Chapter 3*, Unit3 1.

Pietrosiuk, A., Lenherr, E.D., Falk, S., Bonemann, G., Kopp, J., Zentgraf, H., Sinning, I., and Mogk, A. (2011). Molecular basis for the unique role of the AAA+ chaperone ClpV in type VI protein secretion. The Journal of Biological Chemistry *286*, 30010-30021.

Pilatz, S., Breitbach, K., Hein, N., Fehlhaber, B., Schulze, J., Brenneke, B., Eberl, L., and Steinmetz, I. (2006). Identification of *Burkholderia pseudomallei* genes required for the intracellular life cycle and in vivo virulence. Infection and Immunity *74*, 3576-3586.

Pollitzer, R. (1959a). History of the disease. In Cholera, Polilitzer, ed. (Geneva: World Health Organization).

Pollitzer, R., S. Swaroop, and W. Burrows (1959b). Cholera. Monogr Ser World Health Organ *58*, 1001-1019.

Poole, S.J., Diner, E.J., Aoki, S.K., Braaten, B.A., t'Kint de Roodenbeke, C., Low, D.A., and Hayes, C.S. (2011). Identification of functional toxin/immunity genes linked to contact-dependent growth inhibition (CDI) and rearrangement hotspot (Rhs) systems. PLoS Genetics *7*, e1002217.

Pradhan, S., Baidya, A.K., Ghosh, A., Paul, K., and Chowdhury, R. (2010). The El Tor biotype of *Vibrio cholerae* exhibits a growth advantage in the stationary phase in mixed cultures with the classical biotype. Journal of Bacteriology *192*, 955-963.

Pradhan, S., Mallick, S.K., and Chowdhury, R. (2013). *Vibrio cholerae* classical biotype is converted to the viable non-culturable state when cultured with the El Tor biotype. PloS One *8*, e53504.

Prochazkova, K., Shuvalova, L.A., Minasov, G., Voburka, Z., Anderson, W.F., and Satchell, K.J. (2009). Structural and molecular mechanism for autoprocessing of MARTX toxin of *Vibrio cholerae* at multiple sites. The Journal of Biological Chemistry *284*, 26557-26568.

Provenzano, D., Kovac, P., and Wade, W.F. (2006). The ABCs (Antibody, B cells, and Carbohydrate epitopes) of cholera immunity: considerations for an improved vaccine. Microbiology and Immunology *50*, 899-927.

Pugsley, A.P. (1984). The ins and outs of colicins. Part I: Production, and translocation across membranes. Microbiological Sciences *1*, 168-175.

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. Proceedings of the National Academy of Sciences of the United States of America *104*, 15508-15513.

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. Proceedings of the National Academy of Sciences of the United States of America *103*, 1528-1533.

R. Simon, U.P., A. Puhler (1983). A Broad Host Range Mobilization System for In Vivo Genetic Engineering: Transposon Mutagenesis in Gram Negative Bacteria. Nature Biotechnology *1*.

Rashed, S.M., Azman, A.S., Alam, M., Li, S., Sack, D.A., Morris, J.G., Jr., Longini, I., Siddique, A.K., Iqbal, A., Huq, A.*, et al.* (2014). Genetic variation of *Vibrio cholerae* during outbreaks, Bangladesh, 2010-2011. Emerg Infect Dis *20*, 54-60.

Raskin, D.M., Seshadri, R., Pukatzki, S.U., and Mekalanos, J.J. (2006). Bacterial genomics and pathogen evolution. Cell *124*, 703-714.

Rawlings, T.K., Ruiz, G.M., and Colwell, R.R. (2007). Association of *Vibrio cholerae* O1 El Tor and O139 Bengal with the Copepods Acartia tonsa and Eurytemora affinis. Applied and environmental microbiology *73*, 7926-7933.

Rego, A.T., Chandran, V., and Waksman, G. (2010). Two-step and one-step secretion mechanisms in Gram-negative bacteria: contrasting the type IV secretion system and the chaperone-usher pathway of pilus biogenesis. The Biochemical Journal *425*, 475-488.

Reichow, S.L., Korotkov, K.V., Hol, W.G., and Gonen, T. (2010). Structure of the cholera toxin secretion channel in its closed state. Nature Structural & Molecular Biology *17*, 1226-1232.

Riley, M.A. (1998). Molecular mechanisms of bacteriocin evolution. Annual Review of Genetics *32*, 255-278.

Rossi, F., Rizzotti, L., Felis, G.E., and Torriani, S. (2014). Horizontal gene transfer among microorganisms in food: current knowledge and future perspectives. Food Microbiology *42*, 232-243.

Rost, B. (1999). Twilight zone of protein sequence alignments. Protein Engineering *12*, 85-94.

Rowe-Magnus, D.A., and Mazel, D. (1999). Resistance gene capture. Current Opinion in Microbiology *2*, 483-488.

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

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*, 343-347.

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.

Russell, A.B., Peterson, S.B., and Mougous, J.D. (2014a). Type VI secretion system effectors: poisons with a purpose. Nature Reviews Microbiology *12*, 137-148.

Russell, A.B., Singh, P., Brittnacher, M., Bui, N.K., Hood, R.D., Carl, M.A., Agnello, D.M., Schwarz, S., Goodlett, D.R., Vollmer, W.*, et al.* (2012). A widespread bacterial type VI secretion effector superfamily identified using a heuristic approach. Cell Host & Microbe *11*, 538-549.

Russell, A.B., Wexler, A.G., Harding, B.N., Whitney, J.C., Bohn, A.J., Goo, Y.A., Tran, B.Q., Barry, N.A., Zheng, H., Peterson, S.B.*, et al.* (2014b). A type VI secretion-related pathway in Bacteroidetes mediates interbacterial antagonism. Cell Host & Microbe *16*, 227-236.

Sack, D.A., Sack, R.B., Nair, G.B., and Siddique, A.K. (2004). Cholera. Lancet *363*, 223-233.

Safa, A., Nair, G.B., and Kong, R.Y. (2010). Evolution of new variants of *Vibrio cholerae* O1. Trends in Microbiology *18*, 46-54.

Sakazaki, R., and Tamura, K. (1971). Somatic antigen variation in *Vibrio cholerae*. Japanese Journal of Medical Science & Biology *24*, 93-100.

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. Proceedings of the National Academy of Sciences of the United States of America *111*, 9271-9276.

Sample, A.K., Fowler, J.M., and Brubaker, R.R. (1987). Modulation of the low-calcium response in *Yersinia pestis* via plasmid-plasmid interaction. Microbial Pathogenesis *2*, 443-453.

Sandkvist, M., Bagdasarian, M., and Howard, S.P. (2000). Characterization of the multimeric Eps complex required for cholera toxin secretion. International Journal of Medical Microbiology : IJMM *290*, 345-350.

Schell, M.A., Ulrich, R.L., Ribot, W.J., Brueggemann, E.E., Hines, H.B., Chen, D., Lipscomb, L., Kim, H.S., Mrazek, J., Nierman, W.C*., et al.* (2007). Type VI secretion is a major virulence determinant in *Burkholderia mallei*. Molecular Microbiology *64*, 1466-1485.

Schild, S., Tamayo, R., Nelson, E.J., Qadri, F., Calderwood, S.B., and Camilli, A. (2007). Genes induced late in infection increase fitness of *Vibrio cholerae* after release into the environment. Cell Host & Microbe *2*, 264-277.

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

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

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 Pathogens *6*, e1001068.

Seed, K.D., Faruque, S.M., Mekalanos, J.J., Calderwood, S.B., Qadri, F., and Camilli, A. (2012). Phase variable O antigen biosynthetic genes control expression of the major protective antigen and bacteriophage receptor in *Vibrio cholerae* O1. PLoS Pathogens *8*, e1002917.

Seed, K.D., Lazinski, D.W., Calderwood, S.B., and Camilli, A. (2013). A bacteriophage encodes its own CRISPR/Cas adaptive response to evade host innate immunity. Nature *494*, 489-491.

Seed, K.D., Yen, M., Shapiro, B.J., Hilaire, I.J., Charles, R.C., Teng, J.E., Ivers, L.C., Boncy, J., Harris, J.B., and Camilli, A. (2014). Evolutionary consequences of intra-patient phage predation on microbial populations. eLife *3*, e03497.

Seitz, P., and Blokesch, M. (2013). DNA-uptake machinery of naturally competent *Vibrio cholerae*. Proceedings of the National Academy of Sciences of the United States of America *110*, 17987-17992.

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 Genetics *10*, e1004066.

Shao, Y., and Bassler, B.L. (2014). Quorum regulatory small RNAs repress type VI secretion in *Vibrio cholerae*. Molecular Microbiology *92*, 921-930.

Sheahan, K.L., Cordero, C.L., and Satchell, K.J. (2004). Identification of a domain within the multifunctional *Vibrio cholerae* RTX toxin that covalently cross-links actin. Proceedings of the National Academy of Sciences of the United States of America *101*, 9798-9803.

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*, e00106-00111.

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.

Silverman, J.M., Agnello, D.M., Zheng, H., Andrews, B.T., Li, M., Catalano, C.E., Gonen, T., and Mougous, J.D. (2013). Haemolysin coregulated protein is an exported receptor and chaperone of type VI secretion substrates. Molecular Cell *51*, 584-593.

Sixma, T.K., Pronk, S.E., Kalk, K.H., van Zanten, B.A., Berghuis, A.M., and Hol, W.G. (1992). Lactose binding to heat-labile enterotoxin revealed by X-ray crystallography. Nature *355*, 561-564.

Stamatakis, A. (2006). RAxML-VI-HPC: maximum likelihood-based phylogenetic analyses with thousands of taxa and mixed models. Bioinformatics *22*, 2688-2690.

Stamatakis, A. (2014). RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. Bioinformatics *30*, 1312-1313.

Stanley, S.A., Raghavan, S., Hwang, W.W., and Cox, J.S. (2003). Acute infection and macrophage subversion by Mycobacterium tuberculosis require a specialized secretion system. Proceedings of the National Academy of Sciences of the United States of America *100*, 13001-13006.

Stine, O.C., Alam, M., Tang, L., Nair, G.B., Siddique, A.K., Faruque, S.M., Huq, A., Colwell, R., Sack, R.B., and Morris, J.G., Jr. (2008). Seasonal cholera from multiple small outbreaks, rural Bangladesh. Emerging Infectious Diseases *14*, 831-833.

Suckow, G., Seitz, P., and Blokesch, M. (2011). Quorum sensing contributes to natural transformation of *Vibrio cholerae* in a species-specific manner. Journal of Bacteriology *193*, 4914-4924.

Suzuki, Y., and Gojobori, T. (1999). A method for detecting positive selection at single amino acid sites. Molecular Biology and Evolution *16*, 1315-1328.

Swerdlow, D.L., and Ries, A.A. (1993). *Vibrio cholerae* non-O1--the eighth pandemic? Lancet *342*, 382-383.

Takeya, K., Otohuji, T., and Tokiwa, H. (1981). FK phage for differentiating the classical and El T or groups of *Vibrio cholerae*. Journal of Clinical Microbiology *14*, 222-224.

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.

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. Applied and environmental microbiology *56*, 1977-1980.

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. Proceedings of the National Academy of Sciences of the United States of America *84*, 2833-2837.

Teschler, J.K., Zamorano-Sanchez, D., Utada, A.S., Warner, C.J., Wong, G.C., Linington, R.G., and Yildiz, F.H. (2015). Living in the matrix: assembly and control of *Vibrio cholerae* biofilms. Nature reviews Microbiology *13*, 255-268.

Thelin, K.H., and Taylor, R.K. (1996). Toxin-coregulated pilus, but not mannose-sensitive hemagglutinin, is required for colonization by *Vibrio cholerae* O1 El Tor biotype and O139 strains. Infection and immunity *64*, 2853-2856.

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. Infection and Immunity *82*, 1436-1444.

Trokter, M., Felisberto-Rodrigues, C., Christie, P.J., and Waksman, G. (2014). Recent advances in the structural and molecular biology of type IV secretion systems. Current Opinion in Structural Biology *27*, 16-23.

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.

Unterweger, D., Kostiuk, B., Oetjengerdes, R., Wilton, A., Diaz-Satizabal, L., and Pukatzki, S. (accepted). Chimeric adaptor proteins translocate diverse type VI secretion system effectors in *Vibrio cholerae*. The EMBO Journal.

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. Nature Communications *5*, 3549.

VanRheenen, S.M., Dumenil, G., and Isberg, R.R. (2004). IcmF and DotU are required for optimal effector translocation and trafficking of the *Legionella pneumophila* vacuole. Infection and Immunity *72*, 5972-5982.

Veesler, D., and Cambillau, C. (2011). A common evolutionary origin for tailed-bacteriophage functional modules and bacterial machineries. Microbiology and molecular biology reviews : MMBR *75*, 423-433, first page of table of contents.

Villeneuve, S., Boutonnier, A., Mulard, L.A., and Fournier, J.M. (1999). Immunochemical characterization of an Ogawa-Inaba common antigenic determinant of *Vibrio cholerae* O1. Microbiology *145 (Pt 9)*, 2477-2484.

Villeneuve, S., Souchon, H., Riottot, M.M., Mazie, J.C., Lei, P., Glaudemans, C.P., Kovac, P., Fournier, J.M., and Alzari, P.M. (2000). Crystal structure of an anti-carbohydrate antibody directed against *Vibrio cholerae* O1 in complex with antigen: molecular basis for serotype specificity. Proceedings of the National Academy of Sciences of the United States of America *97*, 8433-8438.

Vogt, S.L., Pena-Diaz, J., and Finlay, B.B. (2015). Chemical communication in the gut: Effects of microbiota-generated metabolites on gastrointestinal bacterial pathogens. Anaerobe *34*, 106-115.

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

Waldor, M.K., Tschape, H., and Mekalanos, J.J. (1996). A new type of conjugative transposon encodes resistance to sulfamethoxazole, trimethoprim, and streptomycin in *Vibrio cholerae* O139. Journal of Bacteriology *178*, 4157-4165.

Wang, T., Si, M., Song, Y., Zhu, W., Gao, F., Wang, Y., Zhang, L., Zhang, W., Wei, G., Luo, Z.Q.*, et al.* (2015). Type VI Secretion System Transports Zn2+ to Combat Multiple Stresses and Host Immunity. PLoS Pathogens *11*, e1005020.

Welch, R.A. (2001). RTX toxin structure and function: a story of numerous anomalies and few analogies in toxin biology. Current Topics in Microbiology and Immunology *257*, 85-111.

Whitney, J.C., Beck, C.M., Goo, Y.A., Russell, A.B., Harding, B.N., De Leon, J.A., Cunningham, D.A., Tran, B.Q., Low, D.A., Goodlett, D.R.*, et al.* (2014). Genetically distinct pathways guide effector export through the type VI secretion system. Molecular Microbiology *92*, 529-542.

WHO (2006). Oral Rehydration Salts. *WHO/FCH/CAH/06.1*.

WHO (2008). WHO position paper on Oral Rehydration Salts to reduce mortality form cholera.

WHO (2010). Cholera vaccines: WHO position paper. Weekly epidemiological record *13*, 117-128.

WHO (2014a). Disease Outbreak News.

WHO (2014b). Antimicrobial resistance: global report on surveillance 2014.

WHO (2014c). Cholera. Fact Sheet N°107

Wiersinga, W.J., Currie, B.J., and Peacock, S.J. (2012). Melioidosis. The New England Journal of Medicine *367*, 1035-1044.

Williams, S.G., Varcoe, L.T., Attridge, S.R., and Manning, P.A. (1996). *Vibrio cholerae* Hcp, a secreted protein coregulated with HlyA. Infection and immunity *64*, 283-289.

Wu, Z., Milton, D., Nybom, P., Sjo, A., and Magnusson, K.E. (1996). *Vibrio cholerae* hemagglutinin/protease (HA/protease) causes morphological changes in cultured epithelial cells and perturbs their paracellular barrier function. Microbial Pathogenesis *21*, 111-123.

Xu, H.S., Roberts, N., Singleton, F.L., Attwell, R.W., Grimes, D.J., and Colwell, R.R. (1982). Survival and viability of nonculturableEscherichia coli and*Vibrio cholerae* in the estuarine and marine environment. Microbial Ecology *8*, 313-323.

Zhang, J., Zhang, H., Gao, Z., Hu, H., Dong, C., and Dong, Y.H. (2014). Structural basis for recognition of the type VI spike protein VgrG3 by a cognate immunity protein. FEBS letters *588*, 1891-1898.

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.

Zheng, J., and Leung, K.Y. (2007). Dissection of a type VI secretion system in Edwardsiella tarda. Molecular microbiology *66*, 1192-1206.

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*. Proceedings of the National Academy of Sciences of the United States of America *107*, 21128-21133.

Zhu, J., Miller, M.B., Vance, R.E., Dziejman, M., Bassler, B.L., and Mekalanos, J.J. (2002). Quorum-sensing regulators control virulence gene expression in *Vibrio cholerae*.

Proceedings of the National Academy of Sciences of the United States of America *99*, 3129-3134.

Zinnaka, Y., and Carpenter, C.C., Jr. (1972). An enterotoxin produced by noncholera *Vibrios*. The Johns Hopkins Medical Journal *131*, 403-411.

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. Biochimica et Biophysica Acta *1843*, 1664-1673.

Zoued, A., Durand, E., Bebeacua, C., Brunet, Y.R., Douzi, B., Cambillau, C., Cascales, E., and Journet, L. (2013). TssK is a trimeric cytoplasmic protein interacting with components of both phage-like and membrane anchoring complexes of the type VI secretion system. The Journal of Biological Chemistry *288*, 27031-27041.