Discovery and characterization of small non-coding RNAs in *Vibrio cholerae* that contribute to gene regulation during infection

A thesis submitted by

Evan Bradley

In partial fulfillment of the requirements for the degree of

Doctor of Philosophy

In

Molecular Microbiology

TUFTS UNIVERSITY

Sackler School of Biomedical Sciences

August, 2012

Adviser: Andrew Camilli, Ph. D.

ProQuest LLC.
789 East Eisenhower Parkway
P.O. Box 1346
Ann Arbor, MI 48106 - 1346

Abstract

Small non-coding RNAs (sRNAs) are being increasingly recognized as critical regulators of a wide variety of processes in bacteria. To investigate the contribution of unknown sRNAs to virulence gene regulation in *Vibrio cholerae,* we undertook a screen to identify previously uncharacterized sRNAs under the control of the major virulence gene activator in *V. cholerae*, ToxT. Using a combination of direct sRNA cloning and sequencing together with a genome-wide ToxT *in vitro* binding assay, we identified 18 putative ToxT-regulated sRNAs. Two of these ToxT regulated sRNAs were located within the *Vibrio* Pathogenicity Island-1 (VPI-1), the genetic element that encodes ToxT and the Toxin Co-regulated Pilus (TCP). We verified regulation of these sRNAs by ToxT and showed that deletion of one of them, now designated TarB, caused a variable colonization phenotype when competed against the parental strain in an infant mouse model of *V. cholerae* infection. Infections progressing for 18 hours or less showed the Δ*tarB* strain was out-competed by the wild type strain, while those carried out longer, showed Δ*tarB* out-competing the wild type. Additionally, if inoculated from a resource poor environment the Δ*tarB* strain also showed decreased colonization relative to wild type. Using a bioinformatic approach, we identified that TarB-mediated regulation of the gene *tcpF* was primarily responsible for the TarB mutant's *in vivo* colonization phenotype. Further investigation of genes regulated by TarB using genome-wide transcriptional profiling of a TarB over-expressing strain revealed that TarB also directly regulates genes involved in iron and amino acid uptake. We determined that TarB has a repressive effect on many genes within the VPI-1, but has an activating effect on

tcpP/tcpH, encoding regulators upstream of ToxT. Taken together, the data suggest that TarB plays an important role in regulating virulence and metabolic genes early after *V. cholerae* infection, but that this repressive effect on virulence genes later in infection may lead to reduced replication *in vivo*.

Acknowledgments

I would like to acknowledge my adviser, Andy, who's limitless enthusiasm and optimism for this work and all the science that goes on in his lab is an inspiration, especially in times when protocols aren't working and the data yield conflicting results. I would like to acknowledge postdocs in the lab, Revati Masilliamani and Ayman Ismail, for giving freely of their time and expertise; their input into my science was critical to the progress of this work. I'd like to thank my class and lab mate, EmilyKate McDonough, for her input into my writing and figure design. We wrote a book chapter together and it was a great experience.

Outside the lab, my mother and father, Linda and John Bradley, have always been a great source of confidence and encouragement; they never doubted this day would come, though I sometimes did. I am indebted to my friend and classmate, Nikolai Schnitke. We've been in this together since college and I couldn't have made it this far without his support.

I would finally like to thank all my close friends in Boston, who were there for me in the worst of times, and with whom I spend most of my free time chasing a plastic toy.

Table of Contents

Chapter 2: High throughput screens and discovery sRNAs regulated by ToxT

- Detection of putative ToxT-regulated sRNA transcripts
- Assessment of sRNA mutants in the infant mouse model of intestinal colonization
- ToxT binds in the TarB promoter region
- The *tcpF* mRNA is a target of the TarB sRNA
- TarB's anti-colonization phenotype is time dependent
- Potential other ToxT-regulated sRNAs
- Discussion

Chapter 3: Investigating other upstream factors controlling expression of TarB.

- Introduction
- Upstream factors influencing TarB expression
    - Anaerobic conditions
    - Stationary phase growth
    - LuxR quorum sensing system
- Analysis of *in vivo* expression of TarB using reporter fusion and quantitative RT-PCR
- TarB is likely an Hfq independent sRNA
- Discussion

Chapter 4: Investigation other downstream genes of TarB and possible direct targets

- Introduction
- A genome wide search for other targets of TarB
- Determination of *in vivo* phenotypes of potential TarB targets
- TarB likely directly interacts with the mRNAs of VC1863 and VCA0686
- Discussion

Chapter 5: Discussion and future directions

- sRNAs as members of the ToxR regulons
- Understanding TarB's role during infection
- Future directions
- Conclusion

Chapter 6: Materials and Methods

- Bacterial growth conditions
- Strain construction
- sRNA deep sequencing
- ToxT overexpression and purification
- Affinity purification of ToxT binding sequences
- Mobility shift assays
- AKI induction experiments
- Anaerobic growth
- Pond water incubations
- ToxT induction experiments

- Northern Blots
- qRT-PCR
- Western Blots
- Mouse infections
- Growth curves
- RNA-seq experiments

References

Appendix

- Table 1: Sequencing results, determination of sRNA clusters with relative abundance between ToxT expressing and ToxT ΔHLH expressing libraries
- Table 2: ToxT *in vitro* DNA pulldown results
- Table 3: Genes differentially regulated at 15 minutes of TarB expression
- Table 4: Genes differentially regulated by 4 hours of TarB expression

List of Figures

List of Tables

Introduction

The Organism *Vibrio cholerae*

The bacterial species *Vibrio cholerae* is the causative agent of the disease Asiatic Cholera [1], which is characterized by voluminous secretory diarrhea. *V. cholerae* is a Gram-negative, flagellated bacterium that can be found free living in aquatic environments. *V. cholerae* is a member of the family Vibrionaceae and shares some characteristics with the family Enterobacteriaceae. The bacterium has a single polar, sheathed flagellum that confers the characteristic rapid darting movement of the organism upon viewing with a microscope. *V. cholerae* can be serologically subdivided into groups (serogroups) by their structurally and antigenically diverse O antigen portion of the outer membrane lipopolysaccharide (LPS). Approximately 206 serogroups of *V. cholerae* have been identified to date. However, only the serogroups O1 and O139 are associated with clinical cholera and have pandemic potential [2,3]. Of these two serogroups, only a subset of strains have acquired the key virulence factors, Cholera Toxin (CT) and the Toxin-Co-regulated Pilus (TCP), required to cause cholera. A lysogenic bacteriophage known as CTXΦ contains the genes for CT as well as accessory toxins whose contribution to pathogenesis is less clear [4,5]. While other virulence factors such as TCP are required by the organism to colonize and cause disease, it is the elaboration of CT that is primarily responsible for the voluminous secretory diarrhea that is characteristic of the disease.

Classical and El Tor Biotypes

The world is currently in the 7th recorded world-wide cholera pandemic [4]. The current strain causing infection is of the O1 serogroup, El Tor biotype, which first emerged in the 1960s, replacing the original O1 classical biotype [6]. For reasons that are not clear, the 7th pandemic is currently the longest running and this is hypothesized to be due in part to differences in the El Tor biotype that allow it to persist in the environment longer than the previous classical biotype. There are important differences with respect to virulence between these two biotypes as well. Previous classical strains have been associated with more severe disease than the current El Tor strains, which appear to more frequently cause asymptomatic carriage of the organism [7]. A possible explanation for this difference is lower expression and specific activity of a secreted neuraminidase, NanH, in the El Tor compared to classical biotype strains. NanH is capable of cleaving terminal sugars of glycolipids present in the membranes of intestinal epithelial cells, revealing additional sites for CT binding, and hence more severe disease [8]. Although NanH activity is important in pathogenesis in animal models [9], its possible that its reduced activity contributing to asymptomatic spread of the disease provides more of an advantage to spread of the organism.

The primary genetic differences between classical and El Tor biotypes mostly relate to the presence of a number of new genomic islands in the latter [10]. Some of these appear to be horizontally acquired elements [11] although the exact nature of these elements is not known. Recent investigations into some genes present in these islands has suggested that they may be involved in regulating chemotaxis [12], but mechanistic explanations for how these genes contribute to El Tor's ability to cause asymptomatic

infections or effectively persist in environments that have never had endemic *V. cholerae* before [13], have yet to be determined.

There was likely be some genetic exchange between classical and El Tor strains in recent history as variations in the amino acid sequence of CT and other virulence factors characteristic of classical strains have occasionally been found in El Tor strains isolated from outbreaks in Africa [7]. While there are many differences between these two biotypes, the presence of the quintessential virulence factors and the way they are regulated appear to be similar and in some cases identical [14,15,16,17]. Thus, the results of much of the research conducted *in vivo* (in infection models) to elucidate virulence factors are in general comparable across strains. Nevertheless, it is now thought that classical biotype strains may now be extinct in endemic areas [7], showing the rapid rate at which the population of disease-causing *V. cholerae* in the environment can change.

Cholera the Disease; Epidemiology and Treatment

Cholera is a human disease characterized by voluminous secretory diarrhea and vomiting [2]. As a result, most patients present with symptoms of severe dehydration. Cholera can be rapidly fatal if not treated promptly with intravenous fluids or oral rehydration solution (ORT) [18]. Patients may loose fluid at rates up to 200 ml/kg/hr and as much as 10% of their total body water over the course of the disease [2]. Despite this severe secretory diarrhea, the intestinal epithelium shows minimal pathology and normal physiological transporters of sodium and glucose remain intact [19], allowing for effective absorption of fluids administered orally, so long as they contain the proper electrolytes.

Treatment with antibiotics can shorten the duration of the disease and may reduce fluid loss by up to 50%, but it does not alter the natural course of infection and is secondary to administration of ORT [2]. Many clinical isolates of *V. cholerae* are also resistant to commonly used antibiotics such as tetracycline [2] and these resistances are commonly encoded on mobile genetic elements, allowing resistance to spread rapidly through a population [20]. Hence, treatment of the disease with antibiotics is restricted to severe cases and those in which the resources for providing ORT are limited.

Attempts to develop effective vaccines against *V. cholerae* have been difficult and, in general, do not result in greater then 50% protection for more then 3 years after administration [2,6]. This may be due to the fact that these vaccines, which contain killed whole cells with or without recombinant CT B-subunit, primarily generate an antibody response against the carbohydrate LPS O antigen of *V. cholerae* that does not result in lasting immunity [6,21]. In addition, there are two subtypes (serotypes) of *V. cholerae* O1 LPS (Inaba and Ogawa) that generate antibody responses that are partially specific and thus not completely cross protective with each other and are not cross protective at all to O139 strains [6,21]. With the emergence of the O139 serotype and the possibility of new toxigenic strains arising rapidly by horizontal gene transfer, it may never be possible to formulate a vaccine that will be universally protective [7]. Despite the fact that lasting immunity has not been achieved, use of the current vaccine has been recommended in areas in which active epidemics are occurring, such as the outbreak in Haiti in 2010 [22,23]. Vaccination may also be effective at reducing asymptomatic carriage of *V. cholerae*, which has been documented as a possible mode of transmission [6]. However,

due to the lack of lasting immunity, prevention measures in endemic areas have focused on improved sanitation and methods to prevent spread of *V. cholerae*.

Persistence in the environment

Toxigenic *V. cholerae* is known to be resident in estuary environments and is thought to be a free living organism in those settings (not just in intermediate persisting stage) [24,25]. In this setting, it its thought to be chitinolytic, living off the exoskeletons of copepods or insect egg masses, and to be present within biofilms on surfaces [25,26,27]. In endemic areas, cholera infections tend to peak during warmer months,which may be associated with blooms of plankton and the copepods that feed on them, that could serve as substrates for *V. cholerae* growth. Based on this knowledge, the simple measure of filtering drinking water through a Sari cloth to remove copepods has shown some efficacy in controlling infection, indicating that low cost, simple measures can be effective at reducing disease burden so long as the locations to target are known [2].

Outbreaks are thought to begin by the introduction of the organism into drinking water sources by weather events or human-made changes in the environment [28,29]. Although the organism usually inhabits estuary environments, it is capable of persisting in fresh water and water with growth-limiting concentrations of carbon or nitrogen sources, possibly in a viable but non-culturable state (VBNC). Under such conditions, the organism can cause disease in some animal models of infection and can be detected with molecular methods, but cannot be cultured on standard laboratory media [25,27]. These VBNC *V. cholerae* can be returned to a culturable state by passage in some animal

models [25,27], suggesting that animals could contribute to transition of *V. cholerae* from estuary to drinking water sources, though *V. cholerae* has never been observed to cause disease in animals other then humans. *V. cholerae* present in the aquatic environment is also capable of becoming naturally competent for DNA transformation [30]. In this way, non-toxigenic strains can acquire virulence factors and become novel pathogens to which there is no pre-existing immunity. Alternatively, pathogenic strains can acquire a new LPS O antigen to escape pre-existing immunity, and this is what appears to have occurred with the emergence of the toxigenic O139 serogroup from what is believed to have been an O1 El Tor parent strain [3,31]. This has serious implications for implementation of vaccine or antibiotic treatment regimens as major antigens such as LPS and antibiotic resistance genes could be exchanged in the environment.

The ability of *V. cholerae* to transition from a free-living organism to a parasitic one and back again is of great public health interest because the process could enable the development of cycle-specific targeted measures to prevent spread of *V. cholerae.* A summary of the challenges the organism faces in each environment is shown in Figure 1.1, thus understanding how the organism copes with these different environments will be critical to an effective public health strategy to eliminate the disease. The primary virulence factors of the organism have been known for some time [5,32] and are discussed in the next section. Increasing our knowledge of these factors may allow for more effective vaccines to be developed or targeted treatment of the disease itself that may not involve antibiotics.

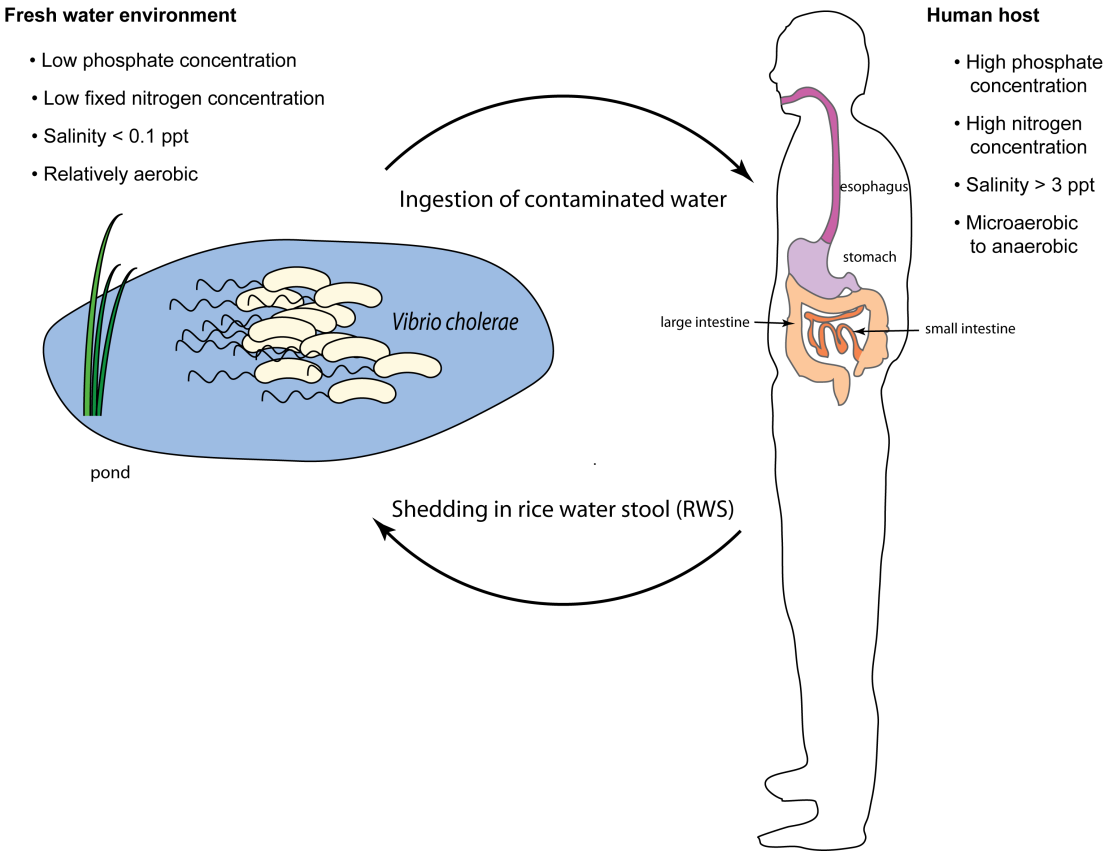

Figure 1.1 A summary of differences in host and aquatic environment

Listed are differences in the freshwater environment and human rice water stool (RWS) that the organism must manage in order to cause a productive infection

Virulence factors of *V. cholerae*

A large number of determinants of *V. cholerae* colonization have been identified in animal models using biochemical and genetic techniques [8,33]. CT was originally identified as the protein in cell-free supernatants that caused watery diarrhea in infant rabbits [34,35]. The TCP was identified as a pilus that was produced coincidently with CT under conditions *in vitro* that stimulate toxin production [8]. The TCP was determined to be a colonization factor by transposon mutagenesis; with transposons containing the alkaline phosphatase gene (TnPhoA); strains with TnPhoA insertions into the major pilin gene, *tcpA*, showed a 100,000-fold drop in LD_{50} in infant mice [32]. Although the CT and TCP represent the primary virulence factors in *V. cholerae*, a large number of accessory factors which aid in motility, attachment, growth, resistance to stresses and other processes inside the host that have smaller contributions to virulence have also been identified [36,37,38].

The TCP is a type 4b, bundle-forming pilus whose major structural subunit is TcpA [39,40]. Type 4 pili are present in a large number of different pathogenic gram negative bacteria and play varied roles in pathogenicity and other cellular processes, such as competence and twitching motility. Most similar to the role of the TCP in *V. cholerae* is the type 4, bundle-forming pilus (BFP) in enteropathogenic *Escherichia coli* (EPEC). The BFP is important for bacterial cell-to-cell attachment and adherence to the intestinal surface in the form of microcolonies [41]. This illustrates the importance of proteins that mediate cell-to-cell adhesion in intestinal pathogens, possibly as a mechanism for resisting shear force in the intestine. In *V. cholerae*, the pilus appears to play a similar

role in micro-colony formation on epithelial surfaces [42]. Similar to the BFP, the TCP is encoded on a mobile genetic element, though essential functions of this element for integration and excision appear to have been lost [43]. The TCP is also the surface receptor for CTXΦ particles [43] that allows binding and infection with CTXΦ particles [5]. While many environmental *V. cholerae* O1 isolates are TCP positive, some do not contain CTXΦ, and thus do not have pathogenic potential. However, laboratory experiments suggest these strains could become lysogenized with the phage [44]. In addition, *V. cholerae* can be transduced with phage during infection of the host intestine [5], which has been used as a tool in the laboratory to investigate TCP expression in animal models of infection [45]. This opens up the possibility that non-toxigenic strains that are positive for the TCP could become converted to toxigenic strains during the infection process, revealing another mechanism by which new variations of pathogenic *V. cholerae* could emerge.

TcpF is another essential factor that is part of the TCP biogenesis operon and requires the pilus for secretion, but is not, strictly speaking, part of the pilus [42,46]. This protein is secreted *in vitro* during virulence factor stimulating conditions, and secretion deficient mutants of the protein do not complement the colonization defect of a Δ*tcpF* mutant *in vivo* [47]. TcpF is also present in TCP positive environmental isolates of *V. cholerae*, however these environmental versions of the protein do not always complement a TcpF null strain *in vivo*, suggesting an ancestral non-pathogenic role for this protein [47]. Although this protein is required for virulence inside the host, it has no noticeable *in vitro* phenotype and its function is unknown.

CT is an AB-type toxin that attaches to intestinal epithelial cells by binding to the

GM1 ganglioside lipid on intestinal epithelial cell surfaces. The A subunit of the toxin is an ADP-ribosylating toxin which targets the G_{α}s-subunit of a stimulatory GTP-binding regulatory protein. This modification results in activation of adenylate cyclase leading to over-production of cAMP [48]. Which results in protein kinase A-mediated phosphorylation of the CTFR [48]. This culminates in protein kinase A-mediated phosphorylation of the CTFR luminal chloride channel that leads to efflux of chloride ions and water into the lumen of the small intestine. It is this biochemical event that is responsible for the voluminous diarrhea that is associated with cholera. Logically, the main function of this toxin seems to be to generate fluid accumulation which the organisms can use to disseminate, and *V. cholerae* deleted for the genes encoding CT do not show colonization defects, at least in the suckling mouse model of colonization [33]. Interestingly, CT also appears to be expressed early in infection in animal models of infection [21,49]; this is presumably when the organism is replicating rapidly, prior to dissemination and release. A possible explanation for this conundrum is discussed below.

Although CT is the primary toxin responsible for the watery diarrhea characteristic of cholera, toxigenic strains also carry accessory toxins which may contribute to disease as revealed by the fact that volunteers inoculated with a Δ*ctxAB* strain of *V. cholerae* still developed diarrhea [8]. These toxins include the hemolysin-cytolysin [50] encoded by the *V. cholerae* genome and the Ace [51] and Zot toxins encoded along with CT on the CTXΦ phage. All of these toxins have been shown to increase fluid accumulation in ligated rabbit ileal loops and disrupt cultured intestinal epithelial layers, although by different mechanisms. The hemolysin-cytoloysin toxin is directly cytolytic and can lead to hemorrhage in addition to fluid accumulation [50]. The

Zot toxin appears to alter intestinal epithelium by disrupting tight junctions (or the Zona Occludens area) of epithelial cells, hence promoting efflux of fluid by disrupting the barriers between cells [52]. The Ace toxin shows homology to eukaryotic ionphores, and may lead directly to efflux of salts from epithelial cells, generating fluid flow in a similar manner to CT [8]. Strains deleted for all of these toxins, including CT produced no adverse reactions in volunteers, however, such strains showed reduced colonization. In conclusion, these accessory toxins do appear to play a role in *V. cholerae* pathogensis and colonization, but the nature of that role is not entirely clear [53].

Natural History of a Cholera infection

The natural history of a cholera infection has been broken down into three phases: initial attachment, replication and toxin elaboration, and detachment and escape. Initial attachment of the organisms to the intestinal epithelium is mediated by the N-acetyl glucosamine binding protein, GbpA [36,54], which also allows the organisms to bind to and grow on chitin *ex vivo*. Once in the small intestine, *in vivo* signals that include bile salts, bicarbonate, temperature and pH lead to activation of the ToxR regulon [55,56]. How these signals result in virulence gene expression is discussed later. The primary virulence factors of *V. cholerae*, CT and TCP, are under the control of this regulon and appear to turned on early in infection [57,58].

Once *V. cholerae* has grown to high density in the small intestine, it begins a down regulation of the primary virulence factors and expression genes necessary for motility and survival outside of the host [49,58]. This has been termed “the mucosal escape response” and is thought to be the genetic program for dissemination from the

host. The mechanism for down regulation of virulence factors remains unclear, but may involve repression of virulence factors by the Pho regulon that is involved in phosphate acquisition [45] and proteolysis of the major positive regulator of virulence, ToxT [59]. This other aspect of the mucosal escape program appears to be under the control of three factors; the stationary phase sigma factor RpoS, the major quorum sensing regulator and transcriptional repressor protein HapR (for hemaglutinin and protease repressor), and the catabolite repression protein CRP [49,60,61]. The primary signals for this response are thought to be the high density of bacteria present in the small intestine and nutrient limitation, possibly of both phosphate and carbon. Genes expressed during this phase of the infection appear to be important for migration of the bacteria away from the intestinal border (motility) and for *ex vivo* survival, as genes for processes such as chitin utilization, biofilm formation and iron uptake are upregulated later in infection [58].

The fluid generated during a *V. cholerae* infection is an isotonic transudate created by the osmotic gradient generated by CT. This fluid visually resembles water after it has been used to cook rice, and is hence commonly called rice water stool or RWS. *V. cholerae* within freshly passed RWS have a different transcriptional profile then those isolated from late in animal models of infection. Most notably, chemotaxis genes are upregulated late in rabbit-ileal loop infections, but repressed in RWS [49,62]. Bacteria in RWS are also hyperinfectious, being about 10-fold more virulent then *in vitro* grown bacteria [62]. The regulators that result in the development of this state and the signals that differ between human and animal models of infection that are responsible for this difference are not yet known, but hyperinfectivity is hypothesized to be a major contributor to epidemic spread of *V. cholerae* [63,64]. Although repression of chemotaxis

genes appears to be a contributor to the hyperinfectious state [65,66], the detailed mechanisms of hyperinfectivity are currently under investigation.

Recently, confocal microscopy has allowed for analysis of subpopulations of bacteria during animal models of infection. This has led to the hypothesis that there is one subpopulation of bacteria that expresses high levels of the primary virulence factors TCP and CT, while another subpopulation prepares for exit from the host by upregulation motility and *ex vivo* survival genes [60]. This may explain some data from population averages showing that CT is expressed during the early stage of infection but not late in the infection in the shed bacteria: one population of bacteria, those bound to the epithelium, may be generating fluid flux from the small intestine to aid in the release of the population that is changing its transcriptional profile to prepare for survival outside of the host. Thus, early during infection, the former subpopulation (high CT and TCP expressors) would dominate, while late in infection, the latter subpopulation (motility and *ex vivo* survival gene expressers) would dominate. It is not yet clear what causes the emergence of subpopulations of bacteria during infection, or what role it has in the infectious process, but it will no doubt change our current models of a human cholera infection.

From the standpoint of prevention of cholera in the future, knowledge about the factors involved and the potential *in vivo* signals that regulate the transition from free-living organism to pathogen could help determine targeted measures to prevent the disease in the first place. A brief review of the current knowledge about these factors and signals as well as gaps in that knowledge is presented in the next section.

Transition of *V. cholerae* from free-living organism to pathogen

The bicistronic operon encoding CT, *ctxAB*, is co-regulated with genes located in the *Vibrio* pathogenicity island-1 (VPI-1). The VPI-1 carries some of the hallmarks of a horizontally transmissible element, and may have been at one time, but reports of its ability to be transmitted as an independent element have not been reproduced [11,43,67]. This co-regulation is carried out primarily by the protein ToxT, a member of the AraC family of transcriptional regulators [68,69]. ToxT activates expression of CT and the toxin co-regulated pilus (*tcp*) operon. Both *toxT* and the *tcp* genes are located within VPI-1. The collection of genes known to be involved in regulating virulence in *V. cholerae* is known as the ToxR regulon [56] and a summary of the factors known to interact with this regulon is shown in Figure 1.3. Upstream of *toxT* are the transcription factor heterodimers ToxR/S and TcpP/H. These complexes reside within the inner membrane. The *toxRS* genes are located in the ancestral *Vibrio* genome and are expressed constitutively [70]. They function to regulate outer membrane protein expression in both *V. cholerae* and non-pathogenic *Vibrio* species [56]. Although *tcpPH* is located within the VPI-1, the Aph proteins and CRP [71], which are located outside the VPI-1, control its transcription. This serves to integrate information about pH [72] and carbon availability, respectively, into control of virulence factor expression, although the reasons for this regulation and their consequences remain unclear.

ToxR/S and TcpP/H directly bind as a complex to the promoter of *toxT* to activate transcription [73,74]. ToxR/S and TcpP/H likely sense some signals in the periplasm and/or inner membrane to initiate binding and transcription from the *toxT* promoter but the nature of this signal *in vivo* is not known [55]. The *in vitro* factors which result in

toxT transcription differ between the classical and El Tor biotype strains and it is not clear how these differences influence infection [55]. Here, I discuss findings relevant to the El Tor biotype as it is the currently circulating pandemic biotype. Factors determined *in vitro*, which contribute to activation of the ToxR regulon include low pH, static growth (possibly by generating microaerobic culture conditions [72,75]), and 37^{o}C temperature [70,76]. The primary result of sensing these signals *in vitro* appears to be initiation of *tcpPH* transcription [14]. Addition of bicarbonate to the culture media also enhances toxin production, but this appears to act at the level of ToxT activity (discussed later), not directly contributing to *tcpPH* transcription [77].

There are also factors upstream of TcpP/H that contribute to virulence gene regulation by repressing production of TcpP/H in response to conditions that may be present *in vivo* or in the environment. Phosphate limitation, which likely occurs in the aquatic environment and possibly late in the course of a cholera infection, prevents transcription of *tcpPH* via the action of the positive regulator of phosphate uptake, PhoB [45]. A high population density of *V. cholerae* as sensed by the Lux and Cqs quorum sensing systems leads to production of HapR that represses production of AphA [61,78,79]. It is possible that dense populations in the form of biofilms occurs in the environment and that inhibition of virulence may be beneficial in that case [80], but this may also function *in vivo* late in infection as discussed above. The CRP protein is also known to interact directly with the *tcpP* promoter by binding to it and preventing binding of AphA in a cAMP-dependent manner [81]. This indicates that carbon-limiting conditions, likely present in the aquatic environment and possibly late in a cholera infection, contributes to the repression of virulence [82].

The bacterial second messenger molecule cyclic di-GMP has also been shown to contribute to virulence gene expression, likely at a level above ToxT production [83], although exactly where this molecule acts is currently unknown. Increased concentrations of cyclic di-GMP in bacteria have been generally associated with a transition from a planktonic, motile lifestyle to a sessile one [84]. In *V. cholerae* specifically, increased cyclic di-GMP concentrations have been correlated with enhanced biofilm formation, inhibited flagellar biosynthesis, inhibited chemotaxis [85], and importantly, inhibition of virulence factors, exactly where this molecule acts is currently uknown although it is hypothesized to occur via proteins that bind cyclic di-GMP known ans plz domain proteins, however, specific mechanisms have remained elusive

Among factors that impact cyclic di-GMP concentrations in cells are quorum sensing via HapR. In addition to affecting transcription of *tcpP/H* via its repression of *aphA*, HapR controls a number of genes that contribute to production of cyclic di-GMP [80]. In addition, AphA controls expression of factors regulating cyclic di-GMP expression, adding another level of regulation to the system [84]. The *in vivo* induced three-component system VieSAB also contributes to the cyclic di-GMP pool, likely during infection [83,86]. However, it should be mentioned that the VieSAB system appears to be more relevant in the classical *V. cholerae* biotype, and what this system senses and how it contributes to the expression of virulence factors *in vivo* is currently under investigation.

In addition to initiating transcription of *toxT*, ToxR/S also represses transcription of *ompT* and enhances transcription of *ompU* [56]. *V. cholerae* expressing OmpT are more susceptible to the anti-microbial action of bile salts, and those expressing OmpU

show enhanced resistance to bile salts as well as other environmental stresses [87,88]. It is thought that this regulation of OMPs may allow specific signals to enter into the cell that then trigger virulence gene induction by stimulating *toxT* transcription through ToxR/S and TcpP/H [87]. However, it was demonstrated that these OMPs are dispensable for colonization, at least in the infant mouse model of *V. cholerae* colonization [88], indicating that they are not directly involved in the pathogenic process nor are they required for virulence gene induction.

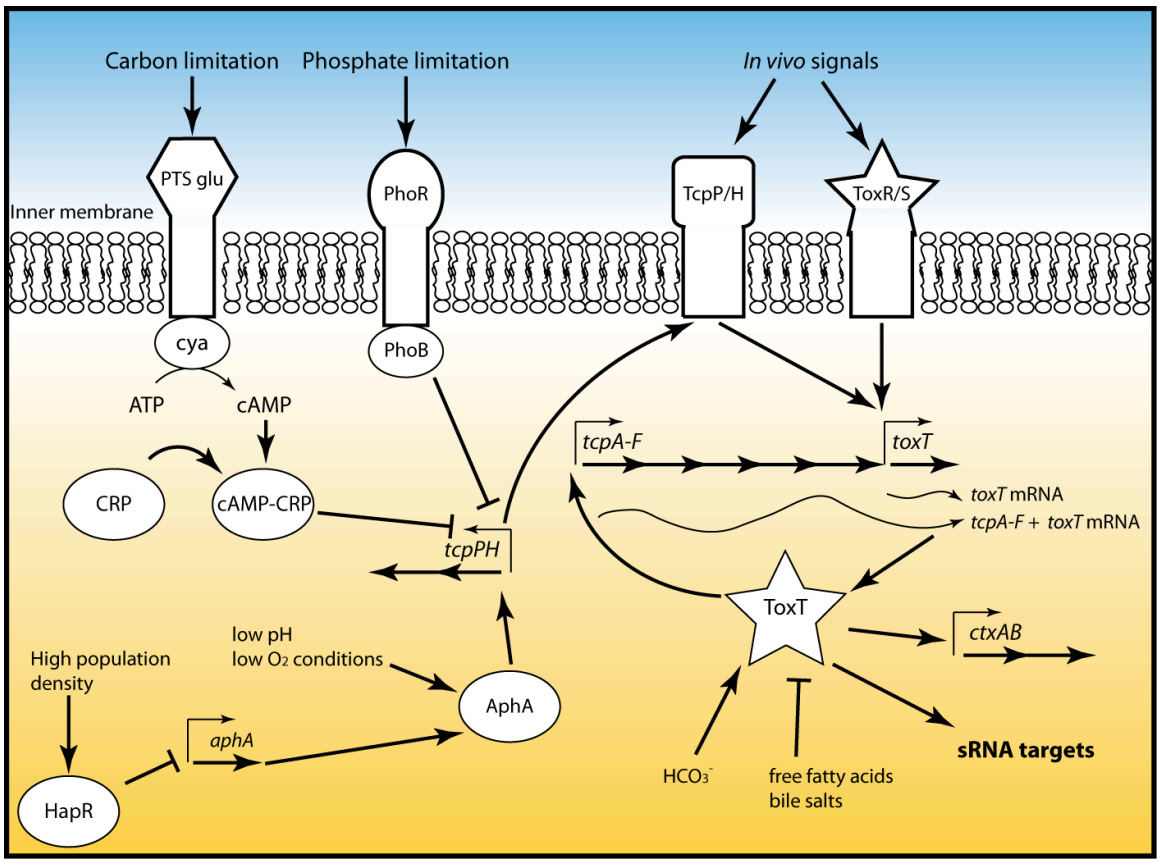

Figure 1.3: Summary of known inputs affecting expression of the ToxR regulons

While the primary *in vivo* signals that activate transcription of *toxT* through TcpP/H and ToxR/S have yet to be determined, many inputs into this regulon have been identified *in vitro*. Both carbon and phosphate limitation have been shown to repress transcription of *tcpP/H* via direct promoter binding of cRP-cAMP and PhoB, respectively. The AphA protein is required for transcription of *tcpP/H* and its activity is enhanced by low pH and low O_2 conditions, but HapR inhibits production of this protein at high cell density. Once produced, ToxT stimulates transcription of the TCP biogenesis operon including the *toxT* gene at the end of the operon, CT genes, as well as sRNA targets, the exact number and nature of which remain to be determined. The activity of the ToxT protein itself can be inhibited by free fatty acids and bile salts and activated by bicarbonate.

Once the ToxT protein is produced, it stimulates transcription of the *ctx* and *tcp* operons by binding to upstream regions within the promoters of these genes, known as *toxboxes* [89], leading to production of *V. cholerae*'s primary virulence factors. The *toxbox* DNA element can be present in different copy number and orientation and thus ToxT displays some flexibility in its ability to bind these different *toxbox* configurations [90,91]. The different known orientations of *toxboxes* as well as a consensus binding sequence determined by two different methods [89,92] are included in Figure 1.4. Although the exact mechanism by which ToxT activates transcription of downstream genes is not known, it is suspected to interact with the alpha subunit of RNA polymerase to recruit it to promoters [93]. ToxT may also relieve repression of virulence gene promoters by displacing the nucleoid protein, Hns [94]. The Hns protein is a DNA binding protein that seems to bind somewhat non-specifically to low GC content regions and repress transcription of a number of genes [95]. Though primarily studied in other organisms, it appears to have an analogous function in *V. cholerae*. The ToxT protein itself appears to also integrate information about the environment into its activity [93]. The protein directly binds unsaturated fatty acids, which inhibit its ability to bind DNA [96]. ToxT activity is also inhibited by temperature and bile salts [93] and activated by the presence of bicarbonate [77], though how these factors interact with the protein directly is not known. Dimerization of ToxT also appears to be critical for its activity [97], which is hypothesized to occur due to the tandem arrangement of most *toxboxes*; however, dimerization is not required for ToxT's ability to bind DNA [98].

ToxT activates additional genes within the VPI-1 by similar mechanisms. These other genes, known as the accessory colonization factor or *acf* genes, do not have known

functions, but were determined by transposon mutagenesis to be important for colonization of the small intestine of infant mice [38,99]. Besides *ctxAB*, ToxT also regulates additional genes outside the VPI-1, including the repression of genes encoding the mannose-sensitive hemagglutanin (MSH) pilus, an anti-colonization factor [98]. In this case, it appears that the binding of ToxT to the *toxboxes* located in the MSH pilus operon interferes with transcription, rather than stimulates it, opening up a new mechanism of ToxT mediated regulation.

The data suggest that regulatory targets of ToxT, be they repressed or activated, are related to the virulent lifestyle of the organism, and by learning what other possible downstream targets of ToxT exist, we can gain insight into how *V. cholerae* causes the disease cholera. While many of the protein coding target genes of ToxT are known [37] or can be inferred from microarrays [60], small non-coding RNA (sRNA) genes that play important regulatory roles, and mechanisms of their regulation by ToxT remain to be investigated fully. These putative additional downstream sRNA genes may be able to teach us more about the environment the organism encounters in the human gut and the processes that are essential for its survival there, which can help target efforts to disrupt those processes and hopefully shorten the natural history of cholera in a patient. A discussion of how sRNAs encoded in bacterial genomes carry out regulation, as well as specific examples in *V. cholerae* and other organisms of how sRNAs contribute to virulence is included below.

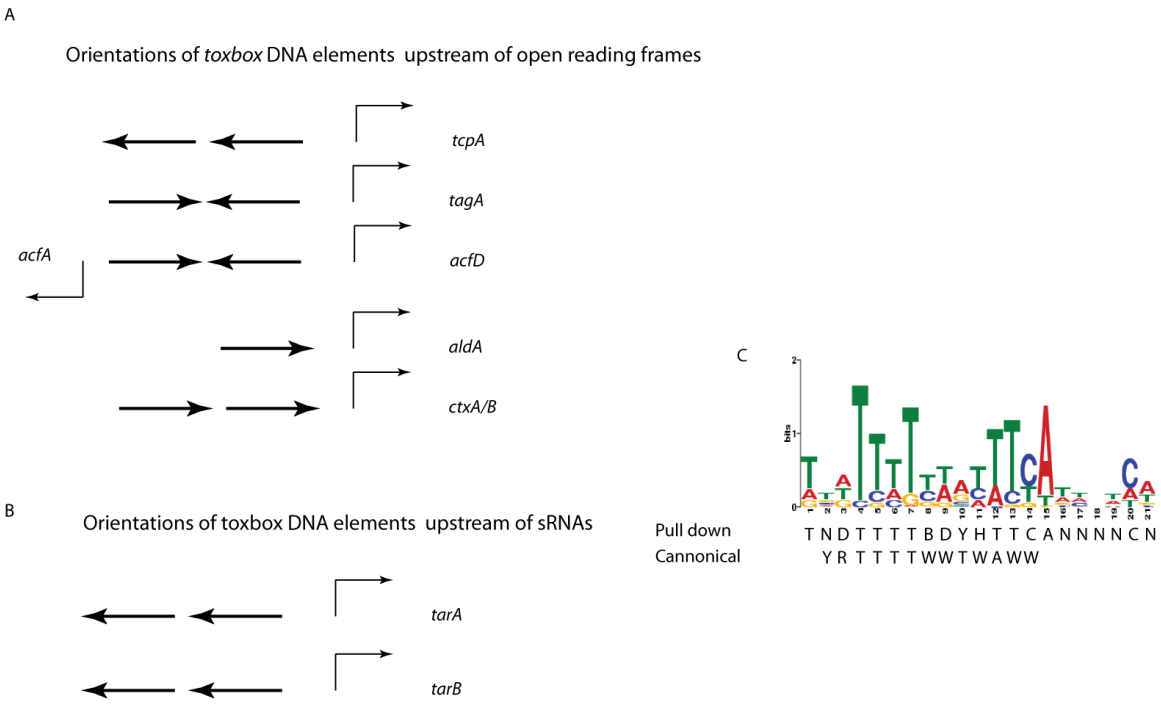

Figure 1.4: Orientation and sequence of *toxbox* ToxT-binding DNA elements

Panel A shows the variety of orientations of ToxT binding sequences relative to the promoters of genes it activates transcription of. Because of the variety of orientations and copy number of these sequences, it is likely that the ToxT protein has some flexibility in its ability to bind DNA and activate transcription of downstream genes. Different orientations of *toxboxes* may relate to different mechanisms of action (e.g. stimulation of transcription as opposed to displacing repressive factors). Panel A shows the ToxT binding sites in the promoters of the two established ToxT-regulated sRNAs. They show similar *toxbox* orientations to *tcpA*, but other ToxT regulated sRNAs with different orientations may exist. Panel C shows the ToxT binding sequence as determined by *in vitro* DNA binding (pulldown) or exhaustive mutagenesis of the *tcpA* promoter (canonical)

Small RNAs in bacterial gene regulation

Recent publications have demonstrated that bacterial genomes encode a large number of sRNAs, that play key roles in regulating a wide variety of cellular processes. Small RNAs were originally discovered in regulation of plasmid genes and first noted in the plasmid ColE1, and as regulators of transposases in several transposons. sRNAs, which are generally between 50-300 nucleotides long, typically do not code for proteins and most often act via base-pairing interactions with their targets [100]. Small RNAs have long been known to regulate plasmid maintenance and toxin-anti-toxin systems [101] but have more recently been shown to be wide-spread and to be key regulators of many important processes in bacteria. These include outer membrane protein (OMP) expression [102], quorum sensing [103] and virulence [104,105]. Their role(s) in many of these processes has been under-appreciated until recently because the genes encoding sRNAs are small enough to be missed by transposon mutagenesis screens, and disruptions by transposons may have been disregarded because of their non-coding nature. However, the development of high throughput massively parallel sequencing systems such as 454 pyrosequencing and the Illumina Genome Analyzer have allowed characterization of sRNAs present in everything from cancer cells to fungi to bacteria.

Many early described sRNAs were determined to be upregluated under stressful conditions, such as when various nutrients are limiting [106] or low pH [107]. As regulators, sRNAs have properties that make them particularly useful under stressful conditions where rapid changes in gene expression are advantageous [106]. Not needing to be translated, they can act more quickly then protein regulators. Additionally, sRNAs

can act on existing mRNAs by either inhibiting or enhancing translation. In this way they can change protein expression profiles without the time it takes to alter transcription or degredation of mRNAs.

The first sRNAs to be characterized were encoded in *cis* and transcribed from the opposite strand as their targets and thus the regulation was carried out by relatively large stretches of perfect complementarity with their targets [108,109]. These *cis*-acting sRNAs most likely regulate a single *cis* target. Many early described sRNAs were determined to be upregulated under stressful conditions, such as nutrient limitation [106] or low pH [107]. As regulators, sRNAs have properties that make them particularly useful as regulators under stressful conditions where rapid changes in gene expression are advantageous [106]. Not needing to be transcribed, then translated, they can act more quickly then protein regulators.

A more recently discovered class of sRNAs act in *trans* and carry out regulation often with imperfect complementarity to their targets [108]. I will focus on this class more extensively. The imperfect complementarity allows for regulation of multiple targets by a single sRNA [110,111]. This class of sRNAs most frequently works in conjunction with the RNA chaperone, Hfq, which can enhance these imperfect base-pairing interactions between sRNAs and mRNA targets [112]. The Hfq protein was originally discovered as a replication factor for the Q-beta RNA phage [113], but it has recently been appreciated that it is involved in most sRNA-mediated gene regulation and RNA metabolism in general.

These trans-acting sRNAs are most frequently described as being transcribed from intergenic regions, although this may be a bias of researchers as quite a few putative

sRNAs appear to be transcribed from areas within open reading frames [114]. Trans-acting sRNAs, generally speaking, have a modular architecture consisting of a 5' seed region important for interacting with targets, an Hfq binding site, and a Rho-independent terminator [115]. A diagram of the typical structure of *trans*-acting sRNAs is shown in Figure 1.2. Studies have shown that these domains can be swapped; the seed region of one sRNA, when replaced with another, can re-direct that sRNA to a new target [116].

Trans acting sRNAs in bacteria have been shown to function via a variety of mechanisms. The most frequently observed mechanism of regulation is via the sRNA duplexing within the 5' untranslated region (UTR) of an mRNA and blocking the ribosomal binding site leading to prevention of translation. In some messages 5' UTRs form structures that occlude ribosome binding and require pairing with sRNAs to disrupt this structure and allow the message to be translated [103,117]. Hence the same basic mechanism of sRNA-mRNA interaction can either lead to inhibition of translation of that mRNA or stimulation of translation. Small RNAs can also pair in other regions of the mRNA. For example, the GadY sRNA of *Salmonella* pairs with the 3' end of a message and enhances its stability [107]. Another possibility is that sRNAs may interact with the coding region of mRNAs. In one documented example of this, an sRNA, which interacted with the coding sequence of an mRNA, enhanced degradation of the mRNA but not its translation efficiency [118]. Other possibilities exist for mechanisms of sRNA-mediated gene regulation and it is clear that some sRNAs can make large contributions to global transcriptional regulation [119]. There may be additional ways in which sRNAs can affect gene expression that have yet to be described.

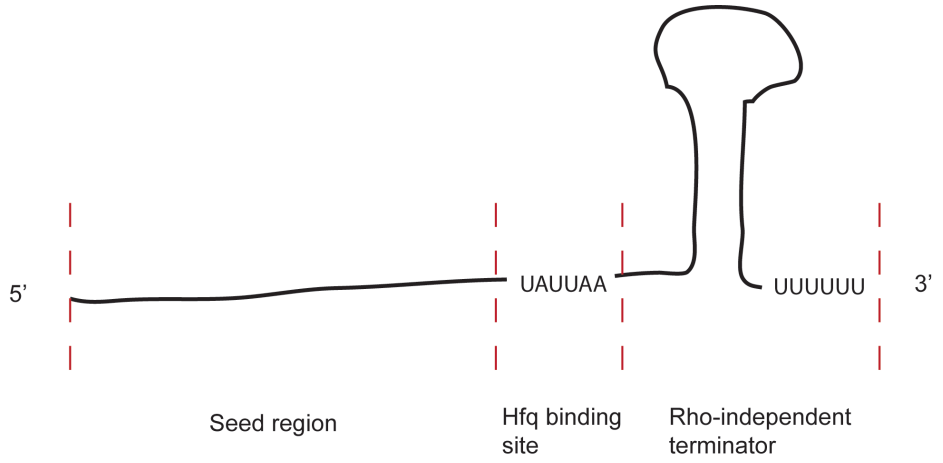

Figure 1.2 Diagrammatic representation of a *trans*-acting sRNA

Shown is the typical architecture of a trans-acting sRNA. The seed region in the 5’ end of the molecule mediates base-pairing and target specificity. 3’ to that is typically a binding site for the RNA chaperone Hfq, this is usually an AU rich sequence. At the 3’ end of the molecule many sRNAs have a Rho-independent terminator that consists of a stem-loop structure followed by a U rich tail.

While direct base-pairing with targets is the most frequent mode of sRNA regulation, some sRNAs act primarily through proteins to mediate regulation. The Carbon Storage Regulator system (Csr) consists of the RNA-binding protein CsrA and three sRNAs that inhibit its function (CsrB-D) [120]. The CsrA protein influences a wide variety of functions in different bacteria [121,122,123] and it does this primarily by binding to a UACARGGAUGU sequence motif on mRNAs and affecting their stability and translation efficiency [120]. The Csr sRNAs contain repeats of this sequence motif to titrate CsrA away from its targets [124].

An example of an sRNA which directly controls translation is the 6s RNA that has been observed to regulate transcription from sigma70 promoters in late stationary phase [125]. The 6s RNA accomplishes this by forming a sequence mimic of sigma70 promoters and competing for binding of RNA polymerase. The Csr sRNAs and 6s RNA exemplify a common theme in which sRNAs serve as sequence mimics of the proteins' other targets. However, this theme is not universal. There is an example of a sRNA that functions as an allosteric regulator of a protein, which is not usually involved in RNA binding. The sRNA Rcd from the ColE1 plasmid, interacts with a cellular tryptophanase to inhibit growth to allow plasmid segregation [126]. It is possible that other sRNAs can interact directly with proteins, which do not usually bind nucleic acids, to modulate there activity.

Much of the regulation carried out by sRNAs via mRNA-sRNA interaction occurs in conjunction with Hfq, a member of the Sm family of RNA binding proteins which includes such diverse homologues as the snRP proteins that carry out splicing in eukaryotes [112,127]. Hfq has homologues that are widely, though not ubiquitously

found in bacterial and archeal species [128]. Hfq is a hexameric protein that binds to AU-rich regions of sRNAs [129] and appears to enhance stability of sRNAs and can facilitate interactions between sRNAs and mRNAs [127,130]. Hfq is central to many regulatory processes involving sRNA-mediated regulation and has been the focus of many sRNA discovery experiments which used immunoprecipitation of Hfq to "pull down" sRNAs for analysis [131,132]. To the best of our knowledge, no attempt to analyze the total number of sRNAs within *V. cholerae* that bind to Hfq has been undertaken, however, there is data that hints at the importance of sRNA-Hfq-mediated regulation to the virulence of *V. cholerae*. Deletion of *hfq* disrupts virulence [133], although it does appear to have pleitropic affects on the physiology of the organism. Because Hfq is central to the activity of many different sRNAs, and since it acts stochiometrically with sRNAs rather then catalytically, it can become a limiting factor when sRNAs are expressed to high levels either under certain stress conditions or artificially by inducible expression systems [134]. This may have implications for sRNA functions physiologically, but certainly has implication for research involving the over-expression of Hfq-dependent sRNAs.

Hfq in conjunction with ribonucleases can contribute to sRNA-mediated regulation. Some sRNAs also require processing by ribonucleases to be active [109,135]. In addition, sRNAs in conjunction with Hfq can stimulate degradation of paired mRNAs by recruitment of poly-adenylation machinery or ribonucleases [109,112]. Research done with an sRNA that inhibits translation by pairing with and blocking the ribosomal binding site of an mRNA suggests that inhibition of translation may be the primary activity of that sRNA and can occur without protein partners [130,136], indicating that recruitment of degradation machinery may be a secondary activity of inhibitory sRNAs.

Pathogens are frequently described as needing to survive stressful conditions related to the natural defenses of the host in order to colonize or cause disease. In this way, it seems logical that the same quick response time afforded by sRNAs would be advantageous during pathogenesis. Certainly *V. cholerae* must survive stressful conditions such as the gastric acid barrier of the stomach to cause infection [8], but what are the direct connections between the virulence regulons and sRNAs? Documented connections between sRNA-mediated gene regulation and virulence are discussed below, although it is clear that we do not currently understand the full picture. Nevertheless, improved tools and sequencing technologies can help to assess the role of sRNAs in the regulation of virulence factors in *V. cholerae* [114].

Small RNA-mediated gene regulation and connections to virulence

It appears that almost any regulon in bacteria, when examined thoroughly, reveals sRNA members, and virulence regulons appear to be no exception. Recent research has revealed sRNA components of the virulence regulons in other organisms, such as *Salmonella enterica* serovar Typhimurium [137], *Listeria monocytogenes* [105] and *Staphylococcus aureus* [137,138,139] (though the *S. aureus* RNAIII, which acts in some ways as a sRNA, is unusually large).

In *Salmonella* many sRNAs involved in virulence and other gene regulation are part of *Salmonella* Pathogenicity Islands (SPIs) that encode systems for bacterial invasion and intracellular persistence [140,141,142]. Small RNAs encoded by the SPI-1 pathogenicity island are involved in both regulating the factors directly responsible for intracellular growth [140] as well as regulating elements of the core genome [142]. In *V.*

cholerae, sRNAs have been shown to be involved in the regulation of virulence factors, but also outer membrane proteins (OMPs) important for the general physiology of the organism [111,143]. Recently, sRNAs that are part of the VPI-1 have been discovered and characterized; one sRNA regulates a glucose-specific phosphotransferase (PTS) [92,144]. This is interesting given that these Pathogenicity islands are thought to be horizontally transmitted elements. Between *V. cholerae* and *Salmonella*, it appears that sRNAs may be a key way that these elements interact with the core genomes of the host. It is also logical that they are involved in the virulent lifestyle of the organisms, as these elements are key to what makes these bacteria pathogens.

In *L. monocytogenes*, an sRNA involved in regulating virulence is also a sensor of S-adenosyl methionine (SAM) [105]. The sRNA, located upstream of the SAM biosynthetic operon, is part of the message that encodes the SAM biosynthetic proteins. In the presence of SAM, this sequence folds to form a transcriptional terminator, leading to the generation of an sRNA that serves to repress expression of the positive regulator of virulence, the transcription factor PrfA. In the absence of SAM, this sequence does not fold to form a terminator but instead allows the expression of SAM biosynthesis genes downstream. Hence, virulence genes become activated in the presence of SAM. This novel dual acting sRNA/Riboswitch opens up new possibilities for sRNA-mediated regulation and shows how sensing of available metabolites may be integrated into virulence gene regulation through the action of sRNAs.

In *S. aureus*, non-coding RNA-mediated regulation of major virulence determinants has been well established for many years in the form of RNA-III. The production of this RNA is controlled by the quorum sensing system that stimulates the

transcription of RNA-III. RNA-III, in addition to encoding a small open reading frame (ORF), acts on mRNAs of *S. aureus* surface proteins and the mRNA of the major transcriptional repressor protein of toxin production, Rot [138,145]. It does this through interactions of the mRNAs of these proteins with unpaired loops in the structure of RNA-III [138]. Although Hfq is implicated in the pathogenesis of *S. aureus*, it does not appear to be critical for this interaction [146]. The *S. aureus hfq* mutant shows a variety of phenotypes, but not gene expression changes related to regulation by RNA-III. Although the *hfq* mutant showed reduced virulence, it is unclear what direct role Hfq-dependent sRNAs have in regulating virulence in *S. aureus* [147].

In *V. cholerae*, there appears to be at least one example of an sRNA having a direct effect on regulation of genes involved in pathogenesis. The sRNA VrrA (*Vibrio* regulator RNA of OmpA) appears to pair with and inhibit translation of *tcpA* mRNA, the major pilin subunit of the TCP [143]. A knock out of this sRNA results in an increased colonization phenotype, presumably because of the release of this negative regulation, but the expression of this sRNA does not appear to be dependent on any members of the ToxR-regulon.

Small RNAs involved in virulence gene regulation do not appear to be unique as far as their mechanisms of regulation, but seem to act in the same ways as previously discovered sRNAs involved in physiologic processes. Thus, general principles learned from these earlier studied sRNAs are applicable to studying this emerging class of sRNAs. In the same vein, the genes that encode sRNAs have many similarities to genes encoding proteins: their promoters and transcriptional terminators share the same elements as those flanking OFRs and therefore likely share similar mechanisms of

transcriptional control [126]. Given this, it has been shown that the same factors, which can upregulate the expression of ORFs, can also activate the transcription of sRNAs. Many sRNAs are upregulated during stress conditions, some of these sRNAs are stimulated by alternative sigma factors, which also activate genes involved in survival in response to stress. The MicA sRNA in *Salmonella* and VrrA sRNA in V. *cholerae* are examples of these [143,148],

Utilization of ToxT to investigate sRNA contributions to gene regulation during infection

Given the existing knowledge about the regulation of sRNA genes, and the fact that ToxT plays a central role in virulence regulation in *V. cholerae*, we hypothesized that we could use ToxT as a tool to investigate sRNAs responsible for virulence gene regulation. ToxT could directly affect the expression of sRNAs by binding in their promoter region and stimulating transcription [149], repressing transcription [98], or leading to anti-repression [150]. Small RNAs regulated by ToxT in any of these ways are likely to be linked to regulation that is important to the infectious process. Determining the targets of this class of sRNAs will therefore provide relevant insight into dynamic gene expression inside the small intestine.

In the transition of *V. cholerae* to a virulent lifestyle from a free living one, sRNAs would have the advantage of being regulators that do not need to be translated before becoming active, so they can act quickly to silence or enhance the expression of existing mRNAs [108,112]. By discovering sRNAs that are controlled by ToxT, and what the targets of those sRNAs might be, we can therefore learn about processes that are

critical for rapid control of gene expression during the disease process. One sRNA regulated by ToxT has already been discovered, which was independently co-discovered by me as described in this thesis, based on the presence of an "orphan" *toxbox* within the *tcp* region [144]. This sRNA negatively regulates a glucose-specific PTS, which has led to the important insight that glucose is likely not a sugar that *V. cholerae* is reliant upon for growth during the infectious process. The discovery of this sRNA opens up the possibility that other sRNAs involved in virulence or physiology during infection may be downstream of ToxT.

Chapter 2. High throughput screens and discovery of sRNAs regulated by ToxT

Portions of this chapter were published as:

Bradley, E. S., K. Bodi, A. M. Ismail & A. Camilli, (2011) A genome-wide approach to discovery of small RNAs involved in regulation of virulence in *Vibrio cholerae*. *PLoS pathogens* **7**: e1002126.

Acknowledgements

Protein purifications were carried out in conjunction with Ayman Ismail who contributed the protocol for generating the tagged proteins and purifying the tagged proteins. Scripts for analyzing the high-throughput sequencing data to determine sRNA abundance and clustering were performed by Kip Bodi. Kip Bodi also performed analysis of ToxT *in vitro* pulldown libraries for enriched sequences. All high-throughput sequencing was performed by the Tufts University Genomics Core Facility. Bharthi Patimalla aided in animal experiments to measure the *in vivo* competitive index of various strains. All other experiments were performed by Evan Bradley.

Detection of putative ToxT-regulated sRNA transcripts

We used sRNA-seq [15], which is a method of direct cloning and deep

sequencing of RNA transcripts 50-250 nucleotides in length [114], to compare a culture in which ToxT or an inactive version missing the helix-loop-helix DNA binding domain (ΔHLH) [57] was expressed from an arabinose inducible promoter on a plasmid (pToxT or pToxTΔHLH, respectively). The highly abundant 5S rRNA and tRNAs present in this size range were depleted prior to sequencing as described [15]. After sequencing we removed residual tRNA and rRNA reads from the data set and aligned the remaining reads to the *V. cholerae* O1 El Tor strain N16916 genome. The number of reads of each unique transcript in each library was normalized to the number of reads of the control, MtlS, an abundant sRNA that controls expression of a mannitol transport system [114] and that does not vary between the conditions tested here (data not shown). A total of 14,578 unique sequences were identified between the two libraries, of which 13,309 were present in only one library or the other. Many sequences not shared between the libraries were very low in abundance and may represent products of random RNA degradation either *in vivo* or during preparation of the libraries. The position of all reads aligned to the N16916 genome and their relative abundances in the two libraries is shown in Table 3. The short sequencing reads were organized into clusters to provide an approximation of each putative sRNA sequence to allow for variations in the start and stop of sequenced sRNAs which may be due to biological variation or variation introduced during library preparation. Many of the 1,269 clusters shared between the libraries had large variations in abundance between the libraries. While this may reflect the true difference in the sRNA transcriptome between these two strains, to help us narrow the list of potential sRNAs we sought a method to determine the subset of sRNAs that was directly regulated by ToxT.

Because sRNA promoters share many characteristics with ORF promoters, it seemed reasonable that any sRNA directly controlled by ToxT would have a ToxT binding site in *cis*. To investigate this we undertook a genome-wide ToxT pulldown of genomic DNA fragments 200-500 bp in length that were modified to allow for subsequent deep sequencing (Figure 2.1 A and 3B), similar to an approach taken with the transcription factor CodY from *S. aureus* [151]. Using a cut off of 3-fold enrichment in pulldown libraries over input libraries, we identified 199 putative binding sites of which 67 overlapped between technical replicates and likely represented the most specific sites (Table S2). A DNA binding motif generated from the 67 enriched sites was a close, though not identical, match to the canonical *toxbox* [89] (figure 2.1, panel C). Of the overall 199 putative binding sites, 64 mapped to the VPI-1, which is consistent with the fact that this locus contains the majority of ToxT-regulated genes. Most, but not all previously described ToxT-binding sites were present in the pulldown library; notably absent were sites within the *tcpA* promoter [89] and sites within the MSH pilus operon [98].

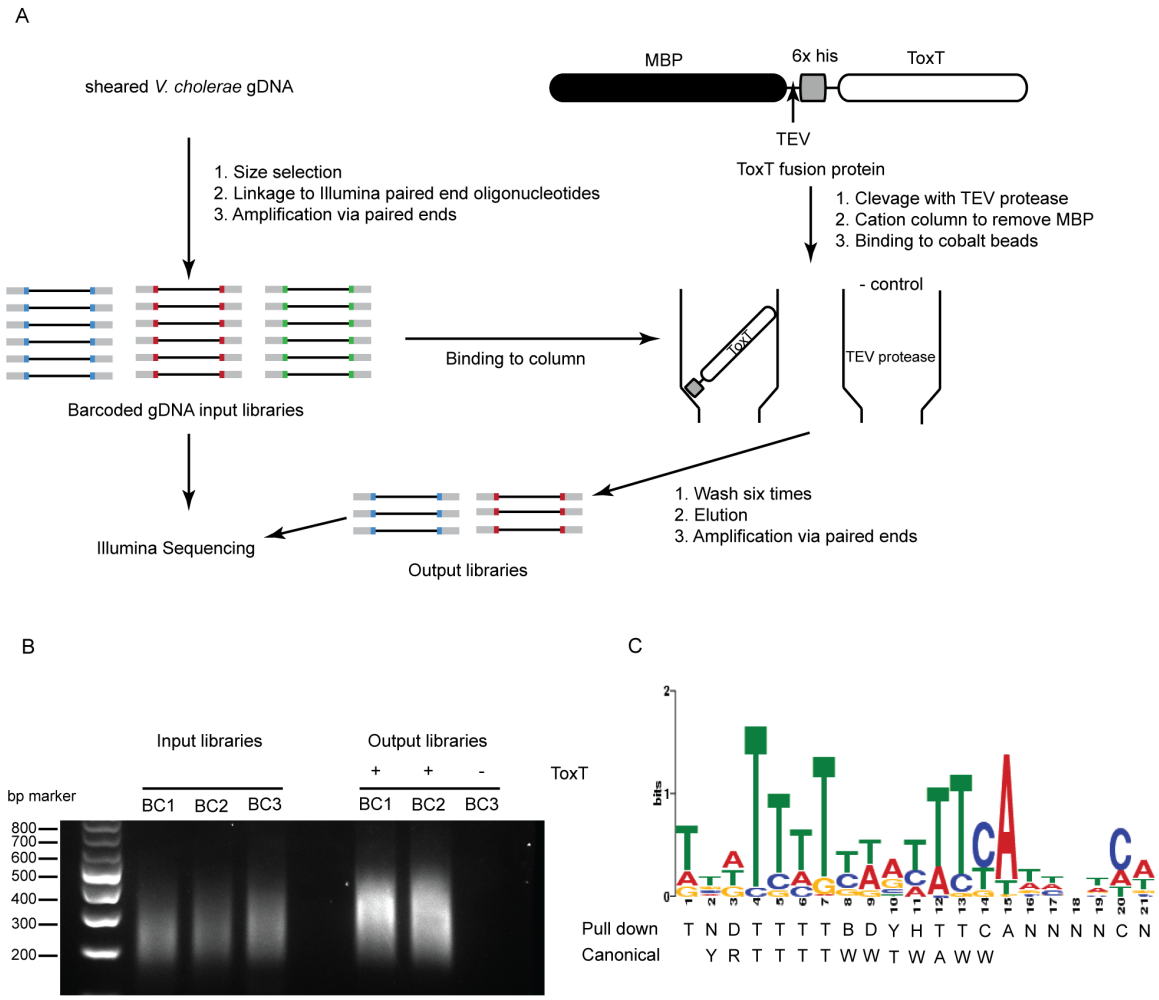

Figure 2.1 Affinity purification of ToxT binding sequences

A) Experimental outline of the ToxT *in vitro* DNA pulldown. Because the purification procedure left a residual amount of TEV protease in the His-ToxT prep, a negative control pulldown was performed with 6His-TEV protease. **B)** Amplification of resulting libraries after pulldowns shows that in the presence of ToxT (libraries BC1 and BC2), DNA was eluted from the column, whereas with TEV bound to the column instead, no DNA is detected after 10 cycles of amplification (BC3). **C)** The resulting binding motif predicted for ToxT present in 66 out of 67 pulldown sites with an E-value of 2.3e-14 as

analyzed by MEME software according to parameters detailed in Methods. The ToxT binding motif predicted by pulldown is shown with the previously reported canonical *toxbox*. Single letter codes are as follows; **B**=C/G/T, **D**=A/G/T, **H**=A/C/T, **N**=A /C/G/T, **R**=A/G, **W**=A/T, **Y**=C/T.

Cross-referencing the putative ToxT binding sites with ToxT-regulated sRNA sequencing data yielded a collection of 18 potential sRNAs transcribed from intergenic regions with *cis* ToxT binding sites. The locations of these pulldown sites, sRNA transcripts and relative abundance between ToxT and ToxTΔHLH expressing strain libraries are shown in Table 1. This analysis revealed two putative sRNAs within intergenic regions in the VPI. To investigate whether these two sRNAs represented genuine transcripts, we probed for each by Northern blot using total RNA from cultures expressing ToxT or ToxTΔHLH. Both of these sRNAs are dramatically upregulated upon expression of ToxT and both are present at the expected size predicted by the sRNA deep sequencing experiment (Figure 2.2). One of these sRNAs was discovered independently by another group and was named TarA [144] (for ToxT activated RNA A). The other, to the best of our knowledge, remains uncharacterized. Since it also showed dramatic up regulation upon expression of ToxT, and given its role in virulence (described below), we named it TarB. Although some upregulation of TarB was seen at later time points in the ΔHLH expressing strain, this is not likely to be due to residual activity of the ΔHLH allele and is most likely due to the culture entering stationary phase as this appears to upregulate TarB independently of ToxT (data shown later). Having now determined that at least two ToxT-regulated sRNAs were present in the VPI, we set out to determine whether they played detectable roles in the virulence of *V. cholerae*.

Table 1. Intergenic sRNAs with *cis* located ToxT binding sites.

Enriched pulldown genome coordinates		sRNA sequencing read genome coordinates[a]		nearby ORFs	ORF annotation; sRNA annotation	Normalized ToxTΔHLH library reads[a]	Normalized ToxT library reads[a]
Start	End	Start	End				
134659	134803	134505	134394	VC0142/ VC0143	hypothetical/ hypothetical	155	375
149092	149248	149280	149445	VC0157/ VC0158	alkaline serine protease/glutamate racemase	0	413
177452	177653	177267	177165	VC0175/ VC0176	transcriptional regulator (putative)/Zinc binding domain protein	242	5630
523047	523177	522904	522819	VC0489/ VC0490	hemolysin (putative)/conserved hypothetical	786	167
889129	889314	888622	888550	VC0825/ VC0826	*tcpI*/*tcpP*; ***tarA***	0	2182
911227	911352	911310	911233	VC0845/ VC0846	putative lipoprotein/phage integrase (degenerative); ***tarB***	2922	519
1037594	1037784	1037758	1037862	VC0971/ VC0972	*ligA* DNA ligase/porin, putative	851	605
1198742	1198847	1199141	1199239	VC1130/ VC1131	*vicH* DNA binding protein/membrane binding protein	798	3

					(putative)		
1412924	1413078	1413070	1413198	VC1328/ VC1329	*D*-galactose or *D*-glucose ABC transporter, permease protein/hypothetical	0	257
2168242	2168432	2168188	2168097	VC2013/ VC2014	PTS system, glucose-specific IIBC component/conserved hypothetical	1078	143
2433764	2433903	2433569	2433387	VC2278/ VC2279	membrane protein, putative/*pepD*	38	10
2549682	2549799	2549443	2549565	VC2384/ VC2385	conserved hypothetical/DNA polymerase	3481	2463
2552739	2553002	2553119	2553020	VC2387/ VC2388	conserved hypothetical/ hypothetical	0	71
180027	180205	182477	182600	VCA0161/ VCA0161	tryptophanase *tnaA*/tryptophan leader peptide *tnaC*	0	198
214601	214731	214400	214531	VCA0197/ VCA0198	GMP reductase (*guaC*)/DNA methyl transferase (putative)	264	205
485328	485491	485507	485625	VCA0546/ VCA0547	conserved hypothetical/ hypothetical	0	525
885950	886158	885901	886039	VCA0932/ VCA0933	hypothetical/cold shock domain family protein	0	3654
886713	886820	893566	893461	VCA0934/ VCA0935	hypothetical/ hypothetical	1134	448

[a]Overlapping clusters hypothesized to represent the same transcript were pooled to determine putative starts and stops and normalized abundances. Normalized abundance scores were rounded to the nearest whole number.

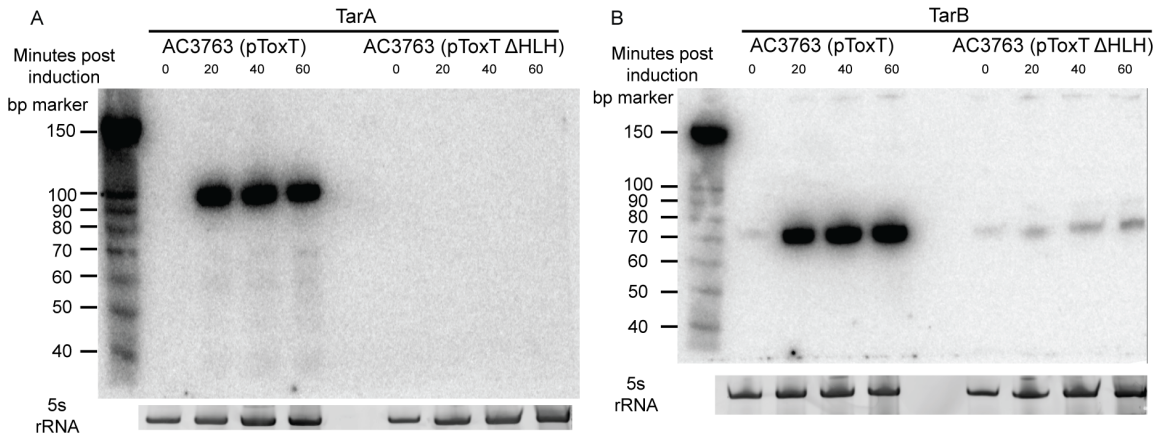

Figure 2.2 Northern blots of TarA and TarB.

^{32}P-UTP labeled riboprobes complementary to sRNAs were used to blot for the presence of the expected sRNAs in total RNA isolated from cultures expressing ToxT or ToxTΔHLH from plasmids. **A**) TarA is detected at the predicted molecular weight and is present at high abundance within 20 minutes after induction by addition of arabinose, which is absent in the transcriptionally inactive ΔHLH form of ToxT. **B**) TarB is also present at the predicted size based on sequencing data and also shows dramatic upregulation in the ToxT expressing strain but not the strain expressing inactive ΔHLH ToxT.

Assessment of sRNA mutants in the infant mouse model of intestinal colonization

Deletion of each sRNA was constructed in the genome and the mutants were tested in completion experimentally with the fully virulent parental strain carrying a Δ*lacZ* marker. No significant difference in virulence was observed for the Δ*tarA* strain either when competed against the parental strain or a strain harboring *tarA* (promoter and *toxboxes* included) on a high-copy vector (Figure 2.3 panel A). It was previously reported that a Δ*tarA* mutant had a decreased fitness relative to its parental strain [144], however, those experiments were performed with a classical biotype strain of *V. cholerae*, and hence regulation by TarA may be less critical or perhaps is masked in the current pandemic El Tor biotype tested here.

In contrast,. the Δ*tarB* strain outcompetes the parental strain by a small but statistically significant factor of 1.6 (Figure 2.3 panel A) suggesting TarB is a negative regulator of virulence. The Δ*tarB* and complemented strains show no change in growth rate or cell yield in Luria-Bertani (LB) broth or in a minimal medium, nor a change in survival in pond water (Figure 2.4).

To see if the negative effect on virulence could be complemented in *trans*, we competed a Δ*tarB* strain containing the sRNA with its own promoter cloned onto a low copy plasmid (p*tarB*) against a Δ*tarB* strain carrying empty vector (pMMB). The Δ*tarB* strain out-competed the complemented strain to an extent that exceeds out competition of the parental strain (Figure 2.3 panel A), which may be due to overexpression of TarB from p*tarB*. If expression of TarB is detrimental to colonization, as these data indicate, the plasmid carrying TarB may be selected against during the infection. To investigate this, small intestine homogenates of infant mice infected with a strain carrying p*tarB*

plasmid the were plated on LB agar and colonies were replica plated onto medium containing ampicillin, which selects for colonies containing the plasmid. Consistent with our hypothesis, the plasmid carrying TarB was lost more frequently than the empty plasmid (Figure 2.3 panel B). This was not the case during growth in LB in the absence of antibiotic selection (data not shown).

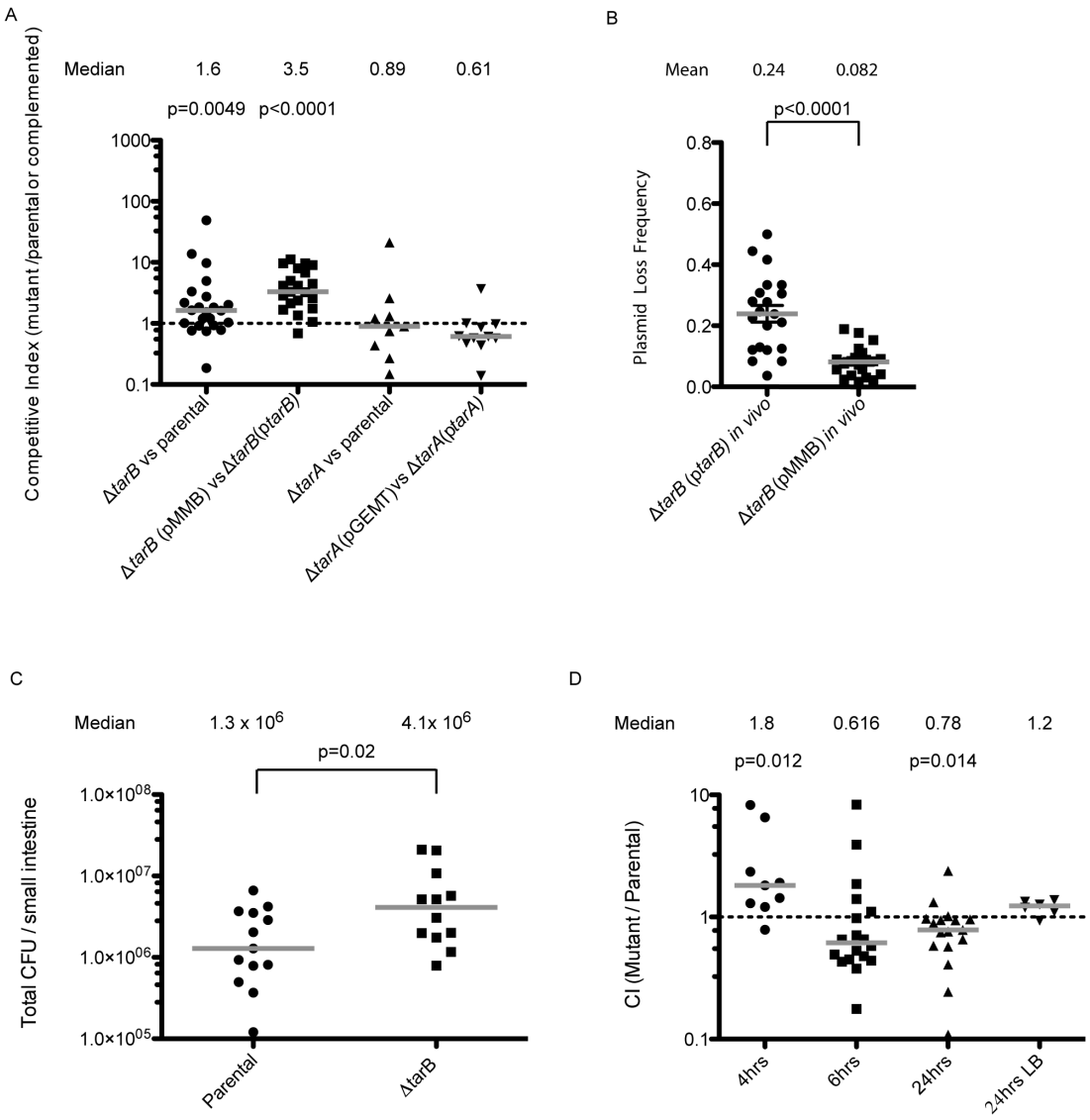

Figure 2.3: Mouse infections performed with ΔsRNA and complemented strains

A) Competitions performed with unmarked deletions of *tarA* and *tarB* against the parental strain carrying a *lacZ* deletion. Competitive indices are reported as the ratio of CFUs in the output adjusted for the input ratio. The Δ*tarB* strain shows enhancement of colonization over the parental strain, but the Δ*tarA* strain shows no significant trend. When *tarA* and *tarB* (promoters included) were cloned into complementation vectors and the complemented strains were competed against deletion strains carrying vector alone, the Δ*tarB* strain shows a more dramatic enhancement of colonization over the complemented strain (median CI = 3.5), while the Δ*tarA* complemented strain shows a slight but not significant advantage over the deletion strain (median CI=0.61; Wilcoxon signed rank test on log transformed data). **B)** Output plates from the competitions in panel A were replica plated onto plates containing ampicillin to assay for presence of the plasmid. Replica plating shows that p*tarB* is lost 2.9x more frequently then the vector alone ($p<0.01$, two sample T-test with Welch's correction for unequal variance). **C)** Single strain infections were performed with wildtype and Δ*tarB* mutants, results are reported as the total CFUs estimated in small intestine homogenates of infected infant mice, in single strain infections the Δ*tarB* mutant also shows an increased colonization phenotype relative to wildtype ($p=0.02$ two sample t-test on log transformed data). **D)** Competitions were carried out in mice after preincubation of the Δ*tarB* and parental strains in pond water for 4, 6 and 24 h. The Δ*tarB* mutant preincubated for 4 h in pond water has a fitness advantage over the parental strain, similar to competitions performed without pond water preincubation. Competitions performed after 6 h of preincubation show no significant trend. However, after 24 h of preincubation there is a reversal of the

above phenotype with the parental strain having a significant advantage over the Δ*tarB* mutant ($p < 0.05$, Wilcoxon signed rank test on log transformed data). Importantly, these strains do not show any difference in fitness during growth in LB after 24 h pond incubation.

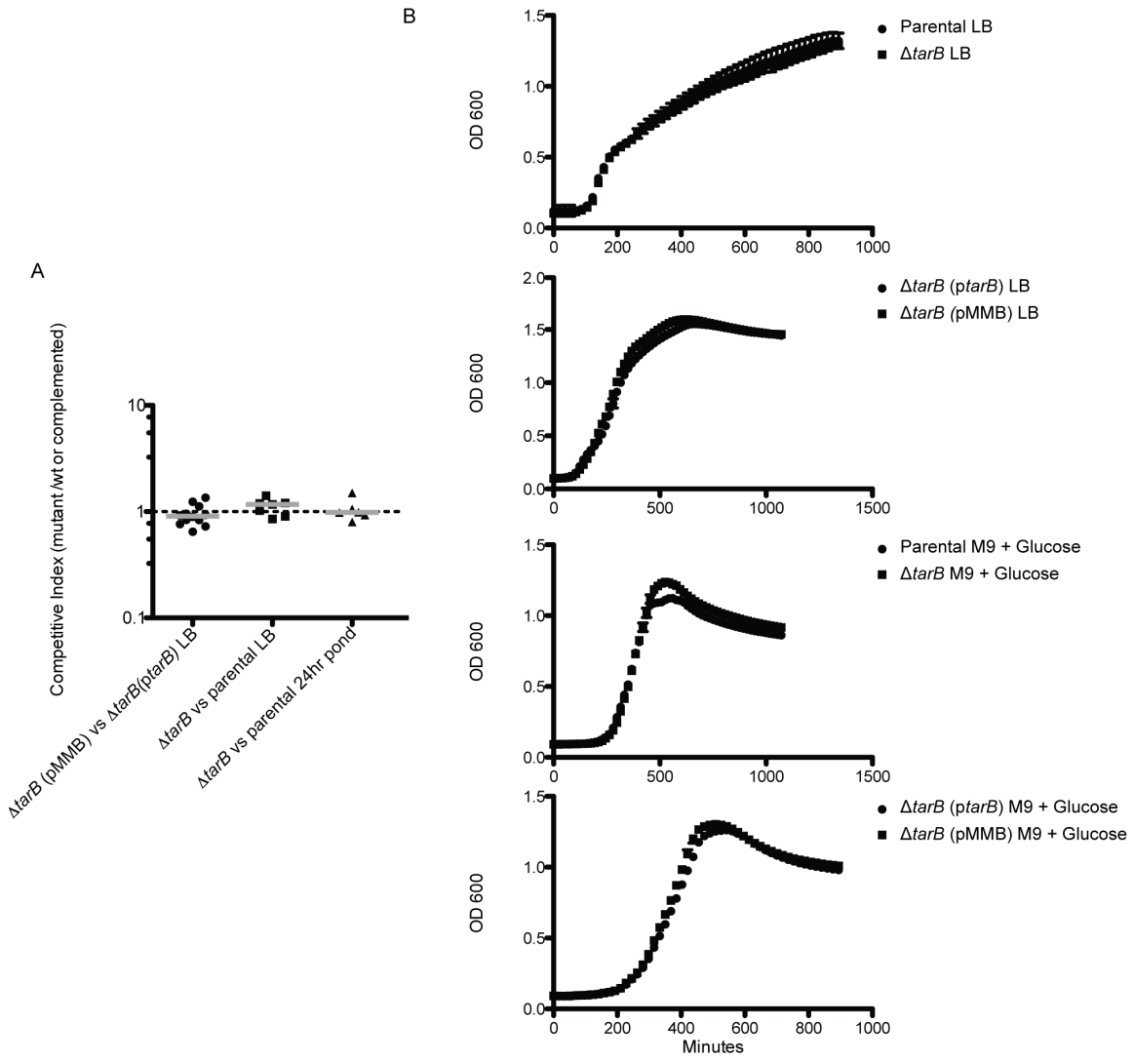

Figure 2.4: *In vitro* analysis of the Δ*tarB* mutant and complemented strains.

A) *In vitro* competitions in LB between the Δ*tarB* and parental strains show no difference in fitness. In addition, the Δ*tarB* strain complemented with p*tarB* or containing empty vector show no significant difference during growth in LB. The Δ*tarB* strain was also

competed against wild type for 24 h in pond water and again the Δ*tarB* strain showed no significant difference in fitness (one sample t-test). **B)** In either LB or in M9 minimal medium with glucose, the Δ*tarB* mutant showed no difference in growth rate when compared to the parental strain. Shown is the median value of growth curves performed in biological triplicate with each individual sample being analyzed in technical triplicate. We also measured the growth rate of complemented strains (Δ*tarB* [pMMB] and Δ*tarB* [p*tarB*]) in both LB and M9 minimal media and these also show no changes in growth rate.

For further confirmation of the hypercolonization phenotype of the Δ*tarB* mutant, we performed single strain infections with the Δ*tarB* and wildtype strains (Figure 2.3C). Total colonization in these two strains indicated that, as seen in competition experiments, the Δ*tarB* mutant showed significant hypercolonization reflected by increased CFUs in the output.

The out-competition phenotype of the Δ*tarB* strain in infant mice and the more drastic attenuated phenotype of the complemented Δ*tarB* strain suggest that TarB is deleterious to colonization of the small intestine. The model that TarB is positively regulated by the master virulence gene activator ToxT, yet functions as a negative regulator of virulence, is counterintuitive. To investigate this model further we performed competitions after incubation of the competing strains for varying times in filter sterilized pond water in an attempt to test the strains in a scenario more similar to a natural infection. After 4 hours of incubation in pond water, the Δ*tarB* mutant retained its ability to outcompete the parental strain, but this phenotype was lost after 6 h of incubation in the pond (Figure 2.3 panel D). After 24 h of pond incubation, the parental now had a statistically significant advantage over the Δ*tarB* mutant when competed *in vivo*, but not when competed for *in vitro* growth in LB.

ToxT binds in the TarB promoter region

The sequence upstream of the predicted TarB start site was investigated and revealed putative -10 and -35 sequences, as well as a direct repeat of putative *toxboxes* (Figure 2.5). The 3' end of TarB determined by deep sequencing corresponded to the poly-U tract of a Rho independent terminator. The *toxboxes* upstream of *tarB* are arranged in similar fashion to those upstream of the virulence gene *tcpA* [152]. To confirm binding of ToxT to this site, a DNA probe consisting of basepairs -100 to +1 relative to the predicted transcription start site was assayed for ToxT binding by gel shift assay. ToxT bound to this region with an affinity within the range of other reported *toxboxes* [149], but not to a non-specific probe of similar length consisting of a PCR product of the 4.5S RNA sequence (Figure 2.5).

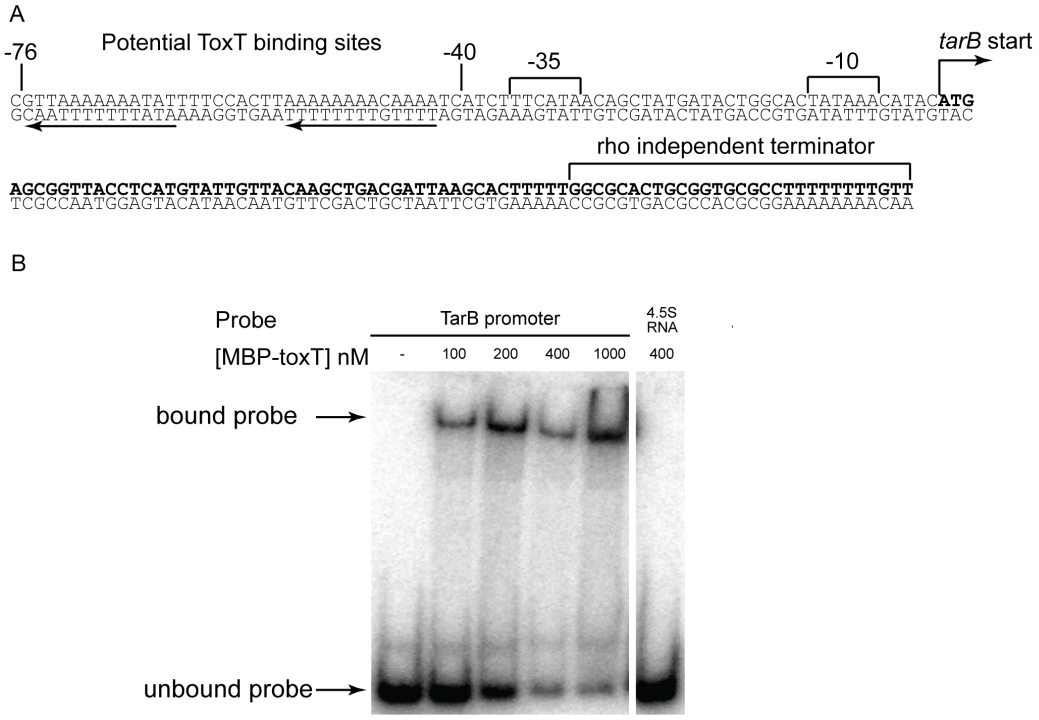

Figure 2.5: Sequence of *tarB* and ToxT binding sites within promoter region.

A) Sequence of *tarB* as determined by the sRNA deep sequencing experiment. Direct repeats of the putative ToxT binding sequence are highlighted by the black arrows. **B)** Electrophoretic mobility shift assays using uncleaved MBP-ToxT fusion protein and the sequence 100 bp upstream of the *tarB* transcriptional start site as a probe. A PCR product of the same size consisting of the sequence of *ffh*, the gene encoding the 4.5S RNA, was used as a negative control probe.

The *tcpF* mRNA is a target of the TarB sRNA

We next wanted to determine the target(s) of TarB that were responsible for the observed negative role of TarB in virulence. Nineteen putative mRNA targets were identified using the program targetRNA [153], which searches for complementarity between the query sRNA and the 5' untranslated region (UTR) of mRNAs of annotated ORFs within a given genome. To validate putative targets, we looked for changes in the steady-state level of the candidate mRNAs using quantitative reverse transcription PCR (qRT-PCR) on total RNA from TarB$^+$ and *ΔtarB* strains both expressing *toxT* from an arabinose inducible plasmid. Of the six putative targets we selected for further analysis only two, *tcpF* and VC2506, had any detectable expression under the conditions tested. When levels of the potential target transcripts were normalized to *toxT* transcript levels, a significant difference between the TarB$^+$ and *ΔtarB* strains was revealed for the *tcpF* mRNA but not for VC2506 (Figure 2.6A). The observed increase in *tcpF* mRNA in the *ΔtarB* background suggests that TarB negatively regulates *tcpF*, which would be consistent with the negative role of TarB in virulence.

To determine if TarB similarly affects TcpF protein expression level, we generated a C-terminal FLAG tag fusion to TcpF in the genome to measure expression by western blot after AKI *in vitro* virulence factor induction [76]. We also generated two sets of three point mutations each within the predicted region of complementarity between TarB and the 5' UTR of *tcpF*, yielding *tcpF** and *tarB** alleles. These mutations are underlined in Figure 2.6B. Because the *tcpF* and *tcpE* ORFs are very close together, there is some overlap between the coding sequence of *tcpE* and the 5' UTR of *tcpF*;

however, the substitutions that were made do not affect the amino acid coding sequence of the upstream gene *tcpE* nor do they alter the Shine-Dalgarno sequence of *tcpF*. Moreover, the mutations were designed to preserve GC content of the region altered. Either set of mutations present alone (*tarB** or *tcpF**) would be predicted to disrupt the interaction between TarB and the *tcpF* 5' UTR while the presence of both is compensatory and would be predicted to restore the interaction.

A strain deleted for *tarB* was then used as the parent strain to construct derivatives having either the *tcpF*-FLAG or *tcpF**-FLAG allele. These two derivatives were then complemented with either p*tarB*, p*tarB** or empty vector (pMMB). These six strains along with the wild type strain carrying the TcpF-FLAG fusion were grown through the static culture phase of an AKI induction and were Western blotted to measure TcpF-FLAG expression. The blots were then stripped and probed for OmpU, an outer membrane protien which is not regulated by ToxT [154], to serve as a loading control. Compared to the wild type strain (Figure 2.6C, first column) the Δ*tarB* and Δ*tarB tcpF** strains carrying the empty vector showed elevated TcpF levels (second and third columns). When the Δ*tarB* and Δ*tarB tcpF** strains were complemented with p*tarB** and p*tarB*, respectively, levels of TcpF remain largely unchanged, indicating that when either the *tcpF* mRNA or *tarB* sRNA are mutated, no interaction can take place and these strains show expression of TcpF similar to the Δ*tarB* mutant. However, when the Δ*tarB* and Δ*tarB tcpF** strains were complemented with p*tarB* and p*tarB**, respectively, to observe affects of the wild type or compensatory interaction when the sRNA is overexpressed, the levels of TcpF drop substantially. Six replicates of this experiment were performed and reveal that statistically significant drops in expression of TcpF occur

only in strains containing either the wildtype TcpF target sequence complemented with wildtype TarB or strains in which the target sequence and sRNA have compensatory mutations (Figure 2.7). When protein samples from these strains were taken after the aeration growth phase of AKI induction and used for Western blots, no differences in TcpF expression were visible (data not shown), which would be expected given the up regulation of *tarB* during the static phase but return to basal level of expression during the aeration phase of AKI induction.

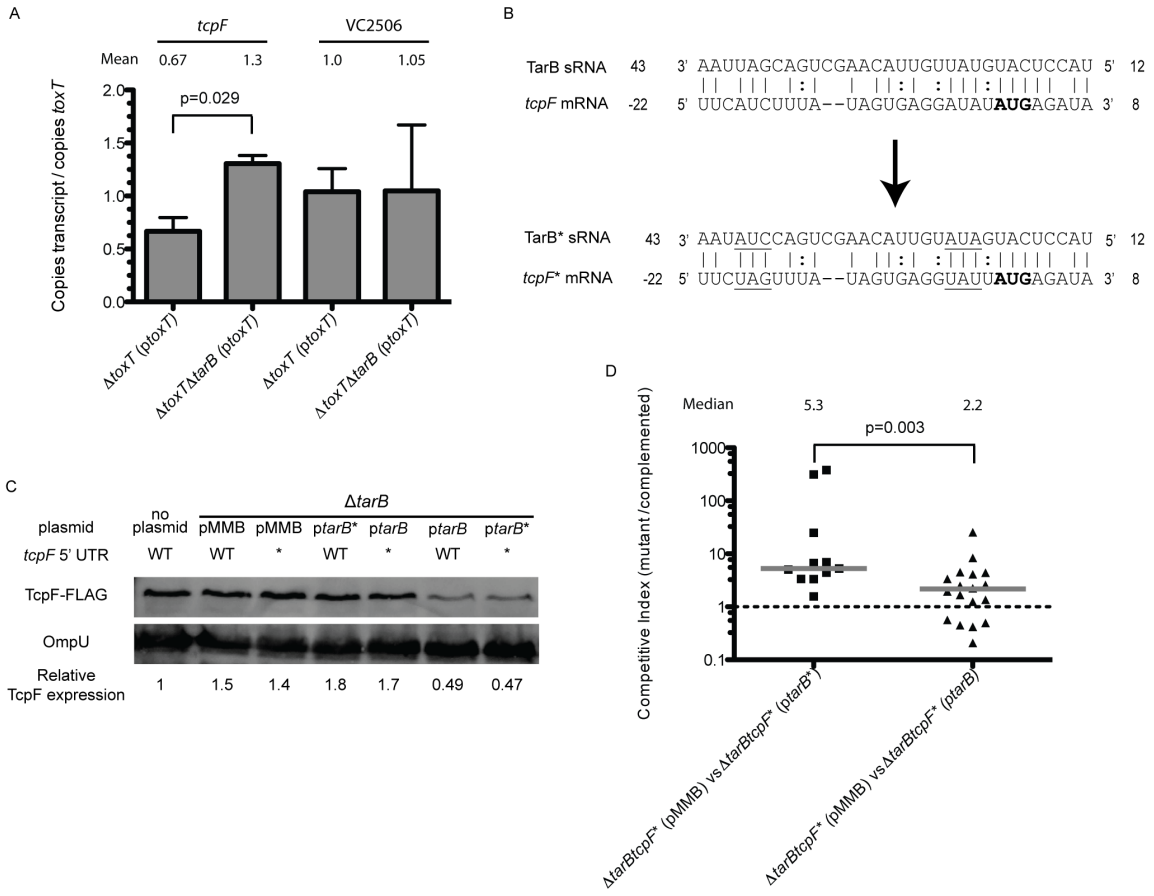

Figure 2.6: TarB interaction with the 5' UTR of *tcpF*.

A) Quantitative reverse transcription PCR was carried out on RNA extracted from strains expressing ToxT from an arabinose inducible promoter. Because the extent of *toxT* induction varied between experiments, transcript levels were normalized to *toxT* transcript. The Δ*tarB* mutant showed a significant enhancement of 2-fold in *tcpF* transcript relative to wild type over the course of four independent experiments (Mann-Whitney U test), the level of the other predicted target (VC2506) did not change. **B)** Predicted base pairing interaction between TarB and the 5' UTR of *tcpF*. The start codon of TcpF is highlighted in bold, the numbering of the *tcpF* transcript is relative to the start of translation, numbering of the TarB transcript is relative to the start of transcription.

The mutations made to generate *tcpF** and *tarB** are underlined. **C)** A fusion of the FLAG peptide to the C-terminus of TcpF was generated to follow TcpF expression by western blot. At the 4 h static time point of AKI induction a band corresponding to the molecular weight of TcpF-FLAG was detected with the anti-FLAG antibody. Blots were then stripped and re-blotted with anti-OmpU antibodies to serve as a loading control. Fluorescence measurements of TcpF-FLAG bands were divided by measurements of OmpU bands. Results are shown for the wild type strain without plasmid (first column) and for the Δ*tarB* strains containing the wild type or mutated TarB cloned on the pMMB plasmid and either the wild type or mutated *tcpF* 5' UTR (*tcpF**) chromosomal allele (columns 2-7). Expression values are standardized to TcpF-FLAG measurements adjusted for loading in the wildtype strain. **D)** Competitions in infant mice between Δ*tarB* strains carrying the *tcpF** allele complemented with p*tarB* or p*tarB** against the same strains carrying pMMB67EH alone. The strain complemented with p*tarB** shows decreased colonization relative to the empty vector strain. When complemented with p*tarB* the competitive index is closer to 1 (Mann-Whitney U test)

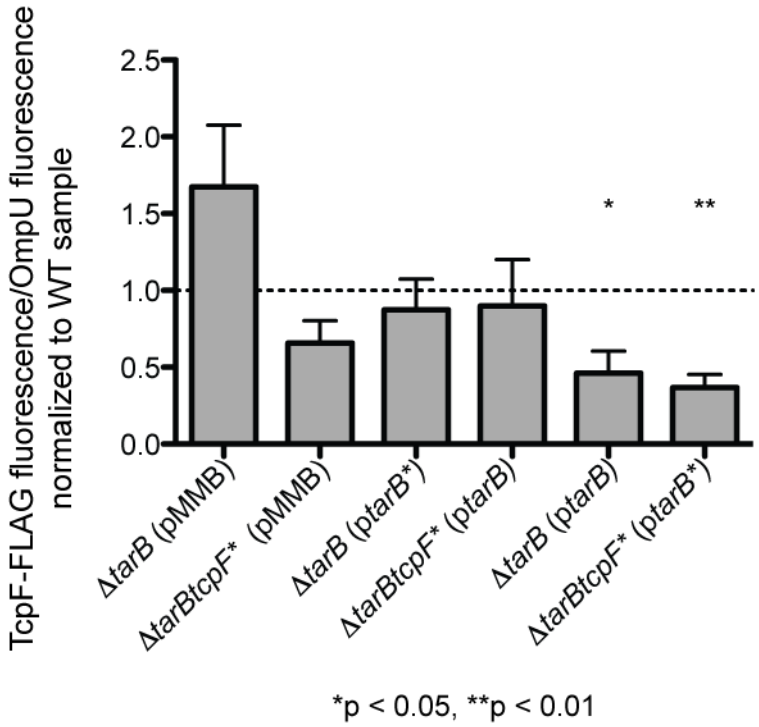

Figure 2.7: Quantitation of TcpF-FLAG Western blot.

Western blotting to quantitate TcpF-FLAG was performed a total of six times for each strain (including the shown example). For each experiment, the TcpF-FLAG fluorescence was divided by OmpU fluorescence and each experimental sample was normalized to the wildtype for that experiment by being set equal to one. Normalized fluorescence values were log transformed and evaluated by one sample T-test against one, the normalized wildtype value. In this analysis only the Δ*tarB* (p*tarB*) strains and Δ*tarBtcpF** (p*tarB**) had mean normalized fluorescence values significantly different from one.

To determine if the interaction of TarB with the 5' UTR of *tcpF* was responsible for the phenotype in mice, competitions were carried out using *tcpF** strain derivatives. We determined that independent of TarB status, that the *tcpF** strains had a colonization defect compared to parental strain (data not shown), this is why we performed all competition experiments in the *tcpF** background to eliminate this as a potential variable. Competition of the Δ*tarB tcpF**(p*tarB**) strain against the same strain carrying empty vector yielded the expected result of out-competition by the latter strain, which lacks *tarB** (Figure 2.6, panel D). Competition of the Δ*tarB tcpF**(p*tarB)* strain against the same strain with vector alone yielded a competitive index that was significantly closer to one, which is expected since neither strain should have an interaction between sRNA and target. The difference between the two competitive indices was highly significant ($p < 0.003$).

TarB's anti-colonization phenotype is time-dependent

Another group has evaluated TarB's colonization phenotype, and they reported a colonization defect in their Δ*tarB* mutant in similar *in vivo* competitions, in contrast to our findings of an increased colonization phenotype [12]. One difference between their assay and ours is the time after inoculation at which they evaluated the competitive index.

We investigated whether the length of infection post inoculation impacted the colonization phenotype of the Δ*tarB* mutant. Consistent with both studies, the colonization phenotype of Δ*tarB* mutant appears to be dependent on the length of time of the competition experiment (Figure 2.8). At earlier time points of infection, the Δ*tarB* mutant displayed a colonization defect. This is particularly interesting as it appears that a undercolonizing strain actually catches up and surpasses the wild type later in infection. The implications of this are unclear, but may relate to an emerging picture that TarB's activity may be most relevant early during infection, whereas the prolonged inhibition of its virulence factors may lead to less replication at later time points, at least in the mouse model of infection. Based on TcpF's hypothesized function of "micro colon maintenance " [42], we can suggest a model that accounts for the observed time-dependent colonization phenotype. Early in infection, TarB's activity appears to be important, most likely prior to the time that TcpF's function is required, hence, misregulation of TcpF may lead to fewer bacteria being able to initially colonize the small intestinal epithelium. As the infection progresses, increased production of TcpF may lead to increased microcolony size, perhaps due to prevention of detachment of bacteria back into the luminal space of the small intestine where it appears that the organisms are not actively dividing [60]. This may account for the time-dependent phenotype that we observe.

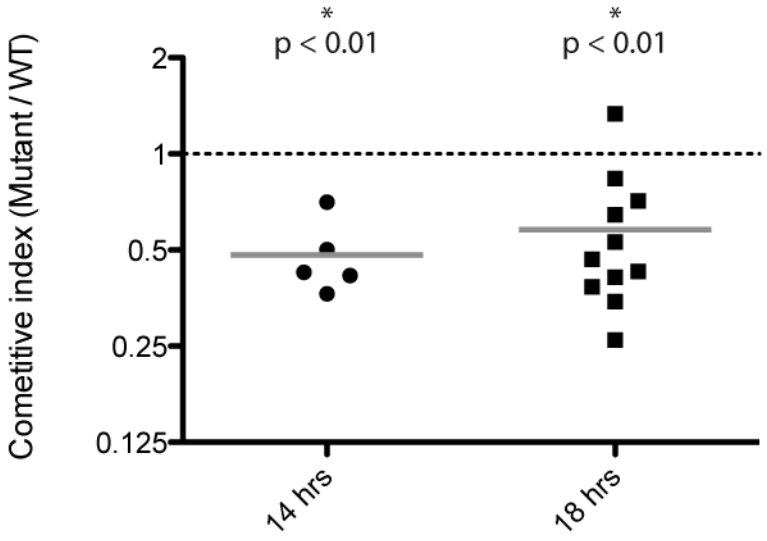

Figure 2.8 The colonization phenotype of the Δ*tarB* mutant is time-dependent Competitions were carried out in infant mice and were allowed to progress for 14 hours and 18 hours as opposed to our previous experiments that were allowed to run for 24 hours. The Δ*tarB* mutant shows a statistically significant colonization defect at 14 and 18 hours ($p < 0.01$ one sample T test)

Other potential sRNA members of the ToxR regulons

Our high-throughput screens identified 18 potential sRNAs that may be regulated by ToxT, and most of these remain to be investigated. Some of the sRNAs were identified in other sRNA discovery experiments [114] or inadvertently, for example, as being regulated by the global iron regulator in *V. cholerae*, Fur [155]. We have performed Northern blots using probes for a number of other potentially ToxT-regulated sRNAs identified by our screen. The sRNA located between the genes VC0175/VC0176 was identified as differentially regulated and having a potential ToxT binding site nearby. Northern blots for this sRNA revealed a larger transcript then expected that was induced upon ToxT expression, though not as robustly as TarB (Figure 2.1, Panel A). We have tentatively named this new sRNA TarC. In addition, this sRNA is induced during the static phase of AKI induction, much like TarB, and this increase is not seen in a Δ*hfq* mutant, though as mentioned in the discussion above, it is not clear whether or not this is due to the requirement Hfq for TarC stability or decreased ToxT induction in this strain. Intriguingly, this sRNA is located just downstream of a previously noted TarB target, VC0177, within the VSP-1 [12]. At this time, It is not clear if there is a connection between these two findings.

Our initial *in vivo* characterization of a mutant made in the TarC putative sRNA *in vivo* has led to some interesting results. The Δ*tarC* mutant shows a small, but statistically significant decrease in its ability to colonize the infant mouse. This growth defect is not observed in competitions carried out in LB broth. We attempted to complement this defect by putting *tarC* on a plasmid, and while the trend in that experiment is towards the

compelemented strain outcompeting the Δ*tarC* strain, the trend was not statistically significant (figure 5.2 panel C). The implications of these findings are not yet clear, but we believe that there is a distinct possibility that other ToxT-regulated sRNAs remain to be discovered and could also yield useful insights into the infectious process.

Although other ChIP-seq style experiments have suggested that ToxT may have no *in vivo* relevant binding sites outside of the VPI-1 or the CTXΦ [12], this is at odds with experiments showing direct binding of ToxT to elements within genes encoding the MSH pilus as a mechanism of ToxT-based repression of those genes [98]. Certainly other *in vivo* factors, such as bicarbonate [77] and fatty acids [96] can influence ToxT's activity and ability to bind DNA, and this may partially account for the disagreement between our results and those from the Mekalanos lab. It is also worth noting that they performed their ChIP experiments after expression of an epitope-tagged ToxT allele from an inducible plasmid during growth in LB, there are two potential issues with this approach. First of all, the physiology of *V. cholerae* grown LB may be dramatically different from the physiology of *V. cholerae* after ingestion by a human from contaminated water, and hence the presence of other factors that can change ToxT's activity may not bet he same in the two environments. Second, the Mekalanos lab used an N-terminally tagged construct of ToxT [12] for their ChIP experiments, and our experience with some N-terminally tagged constructs of ToxT is that they have shown reduced activity relative to the wildtype allele (data no shown) and this may account for their relatively low number of determined binding sites. Hence, there is a possibility that ToxT could bind to and effect expression of genes outside the above mentioned regions, and there is a distinct

possibility based on this work that sRNAs outside the VPI-1 and CTXΦ could be controlled in some way by ToxT.

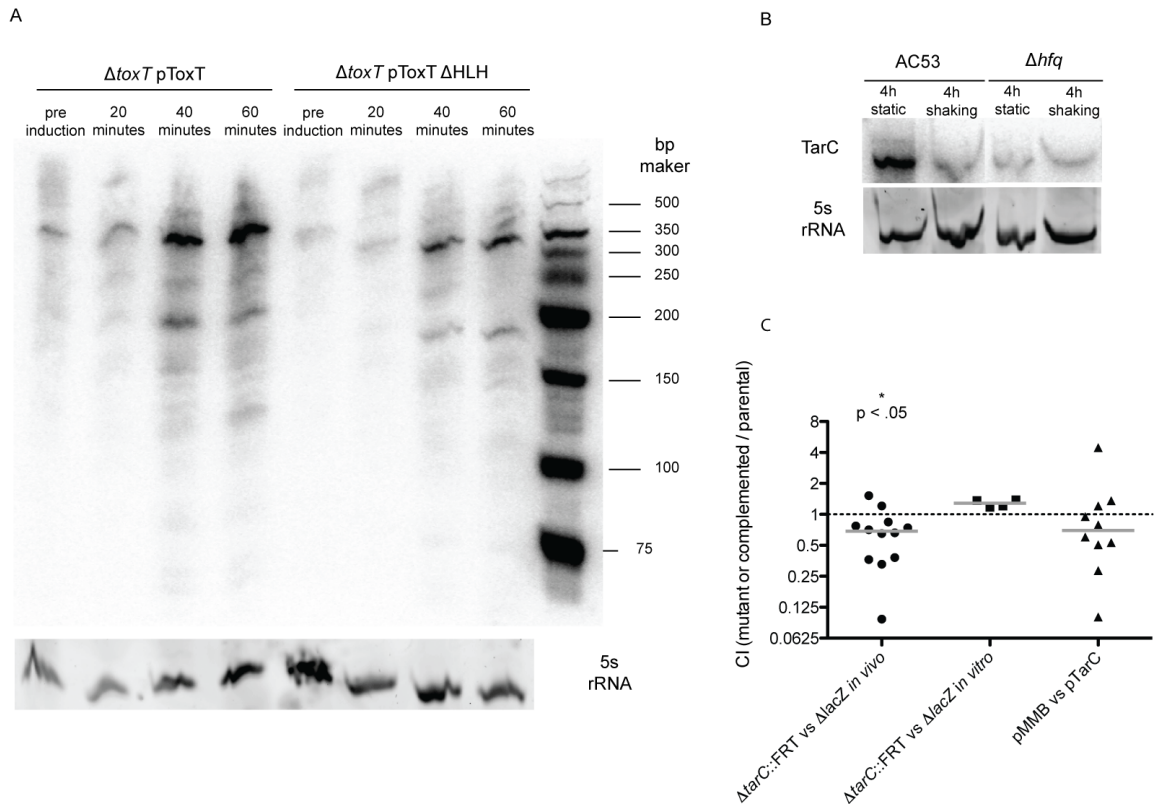

Figure 2.9 A third potential ToxT-regulated sRNA

A) ToxT was expressed from an arabinose inducible plasmid in the Δ*toxT* background. Total RNA was extracted at the indicate time points. At the 40 and 60 minute time points, the band representing TarC is expressed between 3 and 4 fold higher in the pToxT culture vs the ΔpToxT ΔHLH culture. **B)** During AKI induction, TarC shows upregulation during the static phase of culture, similar to TarB. In the Δ*hfq* background, this upregulation is abrogated, most likely due to reduced expression of ToxT under these conditions, as the steady state levels of the sRNA during the shaking phase of AKI induction are similar in the wild type and Δ*hfq* mutant. **C)** When the Δ*tarC*::FRT mutant is competed against a Δ*lacZ* strain, the Δ*tarC*::FRT mutant shows a competitive index of 0.68, ($p < 0.05$

compared to Δ*tarC*::FRT vs Δ*lacZ in vitro*, Mann-Whitney U test). In a complementation experiment, a strain carrying *tarC* (predicted promoter included) on a plasmid showed a trend towards outcompetition of a ΔtarC strain carrying vector alone, although this trend was not statistically significant.

Discussion

Deep sequencing has allowed the interrogation of processes in bacteria with unprecedented detail. Here we used two complementary approaches, deep sequencing of cloned sRNAs and ToxT-bound DNA fragment pulldowns, to identify ToxT-regulated sRNAs. The number of previously estimated ToxT binding sites in the *V. cholerae* genome was between 17 and 20 [89,98]. We have now uncovered what may be a greatly expanded set of targets for ToxT to coordinate expression of protein coding genes as well as sRNAs. The results of the pulldown experiment returned regions of a few hundred basepairs in length that were enriched and many predicted sites are overlapping, which is due to the size range of the fragments used in the pulldown and the automated analysis of the pulldown data. Although many of these sites remain to be validated, we are confident in proposing that the ToxR regulon encompasses more transcripts, both protein coding and otherwise, than was previously thought.

The results of the sRNA deep sequencing reveal the method to be exquisitely sensitive. Because of our exclusion of larger RNA transcripts and depletion of tRNA and 5S RNA in the sRNA size range and the use of Illumina massively parallel sequencing technology, we have achieved tremendous depth of coverage of potential sRNA genes in *V. cholerae* [114]. Transcripts represented by ~40 or more reads could be detected by northern blot (this study and data not shown). However, transcripts represented by fewer than ~40 reads, which may represent low abundance sRNAs, are difficult or impossible to detect by northern blot and other methods such as qRT-PCR are needed for independent validation. Of the 18 candidate ToxT-regulated sRNAs we report here, 11

(including *tarB*) were not identified as putative sRNAs in previous sequencing experiments or bioinformatics-based approaches to sRNA discovery [114,156], displaying the depth of information that can be gained with high throughput sequencing technologies and the conditional expression of sRNAs. In addition to sequencing these transcripts, we have confirmed the existence of 3 individual sRNAs that were detected in the sequencing by sequencing and visualized by northern blot. In comparison to other methods of sRNA discovery, our approach has the advantage of being targeted in its search for ToxT-regulated sRNAs but unbiased in its identification of sRNAs. Approaches utilizing RNA binding proteins such as Hfq [131,132], are not exhaustive as the sRNA we report here likely does not interact with Hfq, though those methods do have the potential to identify mRNA targets as well as sRNAs. Additionally, this approach benefits from the vast strides made in high throughput sequencing recently which generates far more depth of data then microarray based methods [157], including exact 3' and 5' ends and unbiased coverage of positive and negative strand sRNAs. Keeping the latter in mind, this approach can also identify many potential sense and anti-sense sRNAs [114] overlapping with protein coding genes although these potential sRNAs are not discussed here.

In this study we identified a new sRNA member of the ToxR regulon that fine-tunes expression of a virulence factor also within the ToxR regulon, thus adding a new facet to the elaborate virulence gene regulation program in *V. cholerae*. However, when placed in the larger context of *V. cholerae* pathogenesis, it is not entirely clear why a repressor of an essential virulence factor would be produced at the same time as the virulence factor it negatively regulates. The answer may lie in the biphasic nature of *V.*

cholerae gene expression during intestinal colonization [57,60]. The initial induction of virulence factors requires ToxR/S- and TcpP/H-dependent ToxT expression in the intestinal lumen. This is followed by a more robust activation of the TCP and CTX operons closer to the epithelial surface of the small intestine, driven by a positive feedback loop in ToxT expression that is thought to activated in part by the presence of bicarbonate [76,77].

Coordination of TcpF expression by TarB appears to have a positive effect on colonization if the bacteria are coming from a resource poor environment, such as contaminated pond water, or early during an infection from nutrient rich environment. However, even then, the differences in colonization efficiency of the Δ*tarB* mutant are quite small. In contrast, if the infection is allowed to proceed for 24 hours and the bacteria are coming from rich media, overexpression of TcpF in the Δ*tarB* mutant appears to be beneficial. The reasons for this may relate to the details of the experimental system used here, wherein immunologically naïve infant mice are used as a host. In contrast, in nature many hosts in endemic areas will have some level of pre-existing immunity, and may harbor anti-TcpF antibodies as TcpF is a known antigenic protein [42]. It is possible that tight repression of TcpF provides a more pronounced fitness advantage in nature under different conditions then those used here, which would explain TarB's presence among all sequenced isolates of toxigenic *V. cholerae* (data not shown). Further studies into the functional role of TcpF in colonization may shed more light on the necessity of the TarB-mediated post-transcriptional regulation observed here.

Chapter 3. Other factors contributing to TarB expression and regulation

Portions of this chapter were published as:

Bradley, E. S., K. Bodi, A. M. Ismail & A. Camilli, (2011) A genome-wide approach to discovery of small RNAs involved in regulation of virulence in *Vibrio* cholerae. *PLoS pathogens* **7**: e1002126.

Acknowledgements

Neil Greene performed anaerobic growth experiments with help from Brian Mehan. Emily-Kate McDonough provided the *fexA* mutant. Evan Bradley performed all other experiments

Because it is counterintuitive that the same system that activates expression of TarB would also have a negative effect on virulence gene expression, we investigated other factors that could possibly contribute to TarB expression. Expression of virulence factors has been linked to different environmental conditions such as the presence of bicarbonate [77], the presence of free fatty acids [96], and anaerobic conditions [75].

Much of this regulation occurs without necessarily altering expression of key upstream virulence gene activators. Other possible factors responsible for controlling TarB expression could be entry into the stationary phase. Stationary phase regulation of TarB could occur by an alternative sigma factor as was observed for the sRNA VrrA [143], or possibly catabolite repression by the CRP-cAMP complex as carbon sources become depleted [71]. Population density is also integrated into the decision by *V. cholerae* to express virulence factors and the major quorum sensing system acting through HapR can negatively impact expression of virulence factors [79]

Besides investigating other factors that influence TarB expression, we were interested in testing the hypothesis that TarB-mediated regulation coordinates expression of TcpF and other targets spatially and/or temporally during infection. To interrogate expression from the TarB promoter during infection, we constructed a transcriptional fusion of a destabilized (reduced half-life) allele of GFP (GFP-ASV) [60,158] to the TarB promoter. Using this GFP reporter system, we attempted to measure *in vivo* expression of TarB. Unfortunately, we had great difficulty visualizing GFP-expressing bacteria *in vivo* due possibly to low level expression of the reporter. However, as shown in Figure 3.6. we were able to investigate expression of mRNA from the promoter fusion to at least determine at a population average level what the expression of TarB might be during the course of a *V. cholerae* infection with the *in vivo* model.

TarB expression during AKI growth

To further investigate the ability of ToxT to control TarB, we measured

expression of TarB under an *in vitro* virulence factor inducing condition, which is growth for 4 h in static cultures in AKI broth containing sodium bicarbonate, followed by 4 h with aeration [76]. Expression of TarB was induced during the initial static phase of growth, but returned to background levels after 4 h of growth with aeration (shaking) (Figure 3.1A, top panel). The initial induction was dependent on *toxT* as well as *toxR* and *tcpP/H* (Figure 3.1, bottom panel), which are genes upstream in the ToxR regulon that induce ToxT expression [56,159,160]. We also noted that TarB was overexpressed between 7-10 fold in a Δ*tarB* strain complemented with TarB in *trans* despite the fact that *tarB* was cloned under the control of its native promoter. This is, however, consistent with complemented strain's *in vivo* phenotype being more dramatic then the parental strain in competitions with the Δ*tarB* mutant. Some basal expression of TarB was seen during culture in LB, which was greatly enhanced at the transition to stationary phase; however, this increase was independent of ToxT (Figure 3.1B).

During AKI induction in the absence of bicarbonate, ToxT production is stimulated during static growth but the transition to aerated growth is required for CT production [70]. All experiments reported here included bicarbonate in the medium over the course of the experiment, which is sufficient to cause CT production even during growth without aeration [76,161]. In addition, during AKI induction, 4 h of growth in static cultures corresponds with entry into stationary phase [70], which may be linked to expression of TarB as discussed above. Therefore, it is not clear whether the presence of bicarbonate in the media or the growth phase of the culture is the major ToxT-independent contributor to TarB expression during AKI induction.

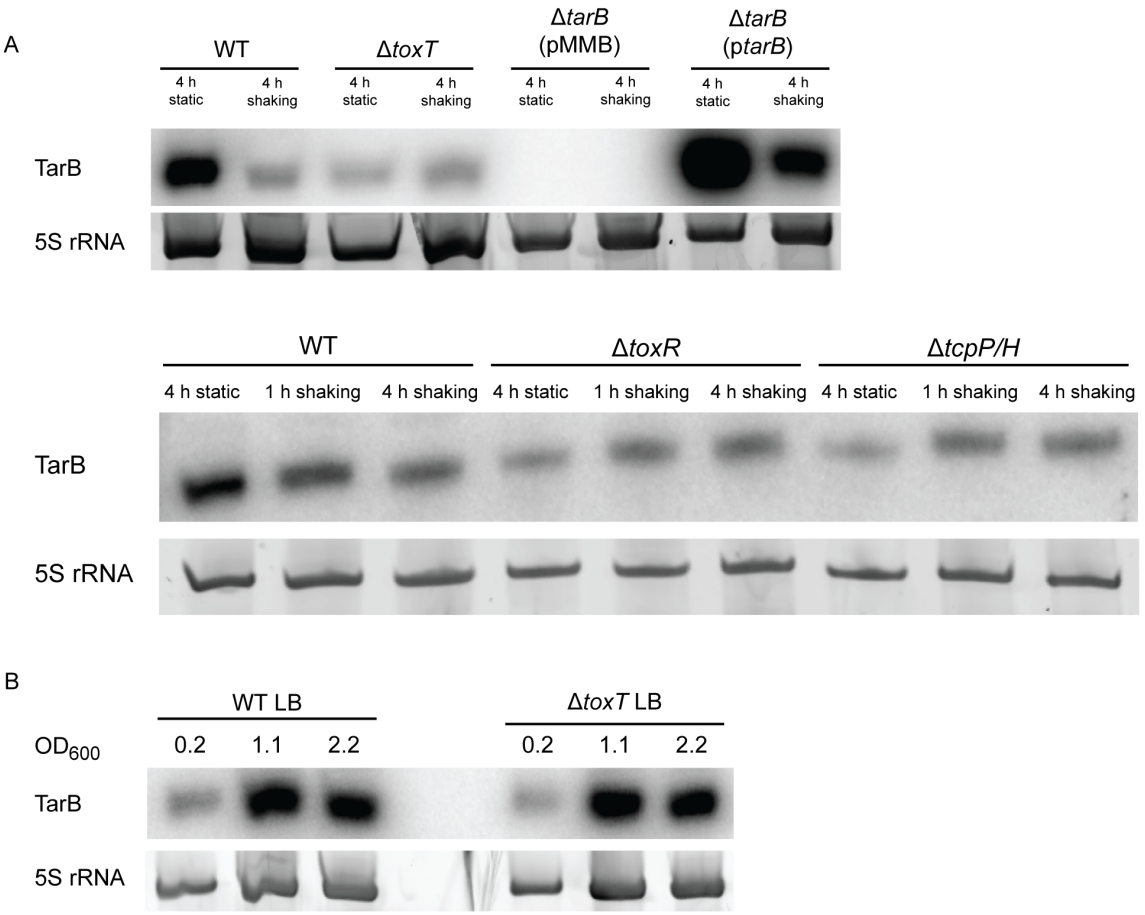

Figure 3.1: Northern blots of TarB during AKI induction and growth in LB.

A) To determine the pattern of TarB expression during virulence factor inducing culture conditions (AKI induction), RNA samples taken after 4 h of static growth, 1 h of shaking growth and 4 h of shaking growth were blotted for the presence of TarB. TarB was upregulated during growth in static cultures (top panels). Upregulation of TarB was also dependent on *toxT, toxR* and *tcpP/H* (bottom panels). Expression from the complementation plasmid p*tarB* reveals that TarB was overexpressed from this plasmid 7-10 fold when adjusting for 5S rRNA loading, though the overall expression pattern of TarB remained the same. The pMMB represents the empty vector negative control. **B)** To

investigate TarB's expression during normal growth in LB, we blotted for the presence of TarB in cultures grown shaking at 37C in both our Δ*toxT* strain and the parental strain. During growth in LB, *tarB* is upregulated upon entry into stationary phase, however, this upregulation was independent of ToxT.

Anaerobic growth conditions contribute to TarB expression

Enhanced expression of TarB during late exponential and stationary phase growth in LB broth and in the static portion of the AKI induction protocol (see above) may be related to oxygen tension in solution. Recent work has shown that AphB directly senses anaerobic conditions via a redox sensitive cysteine residue [162] and is likely what is primarily responsible for driving virulence gene expression under anaerobic condtions. AphB has been shown to be critical for activation of TcpP/H [71,163], which in turn activates ToxT expression. To investigate the contribution of oxygen tension during AKI static growth to TarB expression, we measured expression of *toxT*, *tcpF* and the transcription factor *cadC* by qRT-PCR, and TarB via the TarB-GFP-ASV fusion over the static growth period of AKI induction. The *cadC* gene is activated by the LysR transcriptional repressor homologue protein AphB under low oxygen and low pH conditions [72], and its measurement is used here as a method of determining when the culture is undergoing those conditions.. The results of this experiment are summarized in Figure 3.2. As measured against expression after 2 h of growth under static conditions, expression of *toxT* and *tcpF* more or less reached maximum by 3 h of static culture (Figure 3.2 A, top panels), though expression of TarB-GFP-ASV and *cadC* continued to rise, suggesting additional activation of the *tarB* and *cadC* promoters. Western blotting for TcpF in the TcpF-FLAG fusion strain grown under the same conditions, independently confirmed this finding for TcpF (Figure 3.2A, bottom panels). However, OmpU could not be used as a loading control for this blot as it varies over the course of AKI induction [164], so we did not carry out quantification.

To investigate the contribution of anaerobiosis to expression of TarB, we prepared cultures of wildtype and Δ*toxT* strains in phosphate buffered LB media, to prevent large alterations in pH, and with glucose supplementation to support anaerobic growth [165]. These cultures were prepared in an anaerobic chamber and then grown either aerated in 2 mL or in sealed 10 mL cultures to approximately the same optical density. RNA extracted from these cultures was used in northern blots for TarB (Figure 3.2B). The results showed that anaerobic conditions stimulate TarB expression independently of ToxT. When adjusted for loading, the increases in expression of TarB in the wildtype culture were approximately 2-fold, indicating that under anaerobic conditions, ToxT does drive some expression of TarB. Taken together these results suggest that anaerobic conditions activate TarB. This increase is not likely due to the *V. cholerae* homologue of the anaerobic regulator ArcA (annotated FexA), as a Northern blot for TarB performed on RNA from a *fexA* mutant grown to late log phase showed no changes relative to the parental strain (data not shown). Although this protein has been implicated in regulating virulence genes [166], it does not appear to contribute to TarB regulation under these conditions. The possible contribution of another anaerobic regulator, Fnr, and the possibility of a direct effect of AphB on TarB expression, remains to be investigated

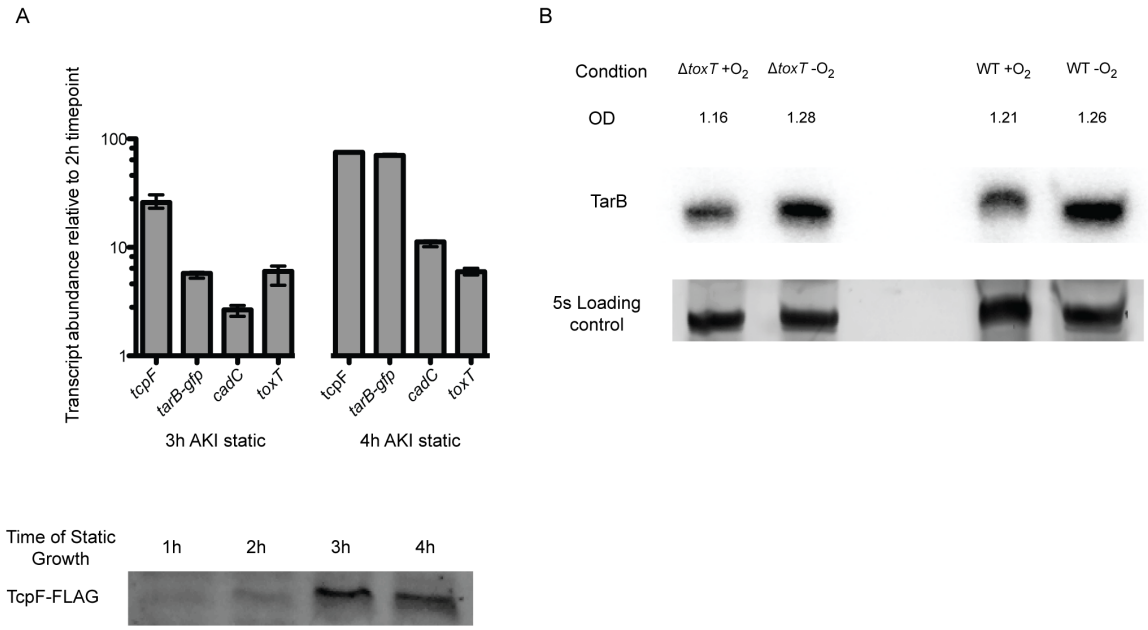

Figure 3.2: TarB expression under anaerobic conditions.

A) Expression of *tcpF, toxT, cadC* and *gfp* from the *tarB-gfp* fusion. The *toxT* and anaerobically upegulated *cadC* genes were followed by qRT-PCR over the course of AKI induction. Shown are median expression values of technical triplicates, adjusted for the *rpoB* loading control relative to the 2 h time point of AKI induction. Results indicate that *toxT* and *tcpF* reached near maximal induction at 3 h of static growth and expression of the *tarB-gfp* fusion and *cadC* showed the most dramatic increases between 3 and 4 h (top panel). This result was confirmed at the protein level by western blot of the wildtype strain carrying the TcpF-FLAG fusion taken through the same AKI induction experiment (lower panel); loading was adjusted for OD as OmpU levels change with activation of ToxR**. B)** Both wildtype and Δ*toxT* strains were grown in buffered media containing glucose either in 2 mL culture tubes with aeration ($+O_2$) or 10 mL sealed culture tubes prepared in an anaerobic chamber ($-O_2$) at 37^{o}C to early stationary phase. RNA was extracted and blotted for TarB. The results indicate that TarB is upregulated under

anaerobic growth conditions independent of *toxT* when adjusting for loading.

Contribution of the quorum sensing cascade to TarB expression

Because quorum sensing and virulence gene expression are closely linked *in V. cholerae* [78], it seemed likely that it may affect TarB expression, possibly through a mechanism independent of HapR's regulation of TcpP/H. The two major quorum sensing systems of *Vibrio* species both converge at the response regulator LuxO, which is phosphorylated and active at low culture densities, but inactive and de-phosphorylated at high cell densities [167]. Phosphorylated LuxO stimulates the expression of a number of sRNAs (known as the qrr or quorum sensing sRNAs) that negatively regulate HapR expression [121]. As previously stated, HapR is inhibitory to the virulence cascade via its inhibition of the *aphA* promoter [78], and thus LuxO provides a link between quorum sensing information and virulence gene activation. Because LuxO controls the expression of a number of other sRNAs and may provide a ToxT-independent means of controlling TarB expression at low culture densities (possibly when TarB's activity is most important), we sought to investigate if the quorum sensing cascade acting through LuxO contributes to TarB expression independently of its effect on the ToxR regulons. We performed northern blots on Δ*luxO*, and a *luxO* constitutively active mutant (LuxO L104D) [168] at various points in the growth curve as well as under virulence factor inducing conditions.

During normal growth, it appeared as though the Δ*luxO* mutant has an increased abundance of TarB, especially early in the growth phase. This increased expression was not seen at later time points in the growth curve, at which point the mutant had the same TarB abundance as wildtype (Figure 3.3 panel A). This pattern fits with LuxO

phosphorylation and activity at low cell densities [169] and implicates LuxO as a repressor of TarB. However, when these LuxO mutants were investigated during *in vitro* virulence factor induction, the opposite results are seen; that is, strains expressing LuxO show increased expression of TarB (Figure 3.3 panel B).

Because LuxO controls expression of the qrr sRNAs that inhibit HapR, that is in turn inhibitory towards expression of TcpP/H upstream of ToxT [78], its expression is critical to induction of ToxT during in vitro virulence factor induction. The results of the experiment shown in Figure 3.3 panel B suggest that LuxO's function in this context (via activation of ToxT expression) contributes more to TarB expression then the mild repression we observed during growth in LB (Figure 3.3 panel A) Although it is unclear what these results indicate, it maybe that LuxO has some function in repressing TarB during normal growth, but once ToxT is induced (something LuxO function is important in allowing), it appears as though that repression is over-ridden.

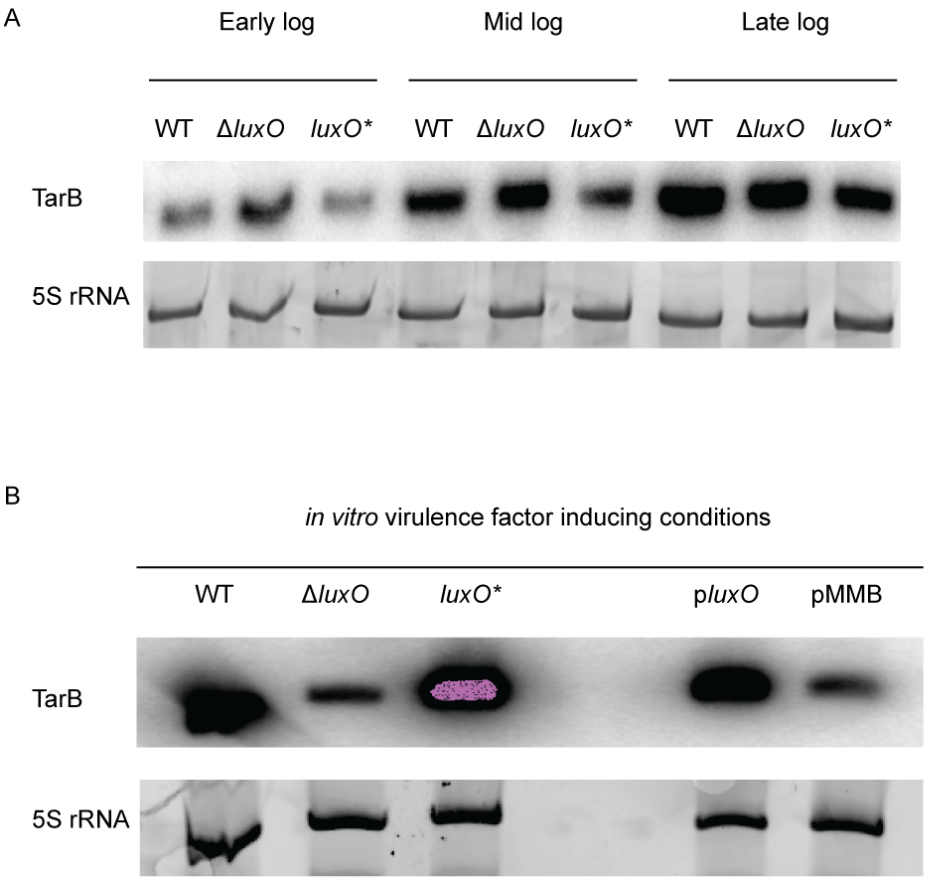

Figure 3.3: LuxO may contribute to the regulation of TarB under normal growth conditions

A) Northern blot for TarB in wildtype (WT), Δ*luxO* and a *luxO* constitutively active mutant during normal growth in LB. Abundance of TarB appears to be higher in the Δ*luxO* mutant early in the growth phase. **B)** Northern blot for TarB in wildtype (WT), Δ*luxO* and a *luxO* and a complemented strain grown under virulence factor inducing conditions. Under these conditions, it appears that LuxO activity contributes to TarB expression, rather then represses it. This is consistent with LuxO acting through the qrr sRNAs to repress HapR and allow expression of virulence factors.

TarB is likely an Hfq-independent sRNA

We also investigated the role of the RNA chaperone, Hfq, in TarB stability and action as many sRNAs that act in conjunction with Hfq are destabilized in its absence [112,170]. The sequence of the TarB sRNA does not reveal a canonical Hfq binding site (UAUUAA) [171] and thus an interaction with Hfq may not occur. The expression of the *tarB* promoter-GFP-ASV fusion was used to measure activity of the TarB promoter during induction of ToxT from the pToxT plasmid in both Hfq^+ and Hfq^- strains. In these same strains, steady state levels of TarB from a native copy of the gene were measured by northern blot. The results of these experiments are summarized in Figure 3.4 and indicate that Hfq does not play a detectable role in stabilizing TarB. In addition, we examined the effect of Hfq deletion on TcpF regulation by qRT-PCR. As shown in Figure 3.4 Panel A, TcpF transcript does not vary greatly between the Hfq+ and Hfq- strains used in the experiment, suggesting that Hfq does not affect the ability of TarB to regulate its targets.

Recent experiments with Hfq have revealed that overexpression of Hfq-binding sRNAs leads to destabilization of other Hfq-dependent sRNAs within the cell, as the over expressed sRNA sequester most of the Hfq present within the cell [134]. To further investigate the Hfq independence of TarB, we examined the effect of TarB overexpression on a known Hfq-dependent sRNA, VrrA [111]. If TarB binds Hfq, then we might expect TarB over expression to sequester Hfq, resulting in the destabilization of VrrA. However, the steady state level of VrrA was not affected by TarB overexpression,

consistent with our hypothesis that TarB is an Hfq-independent sRNA.

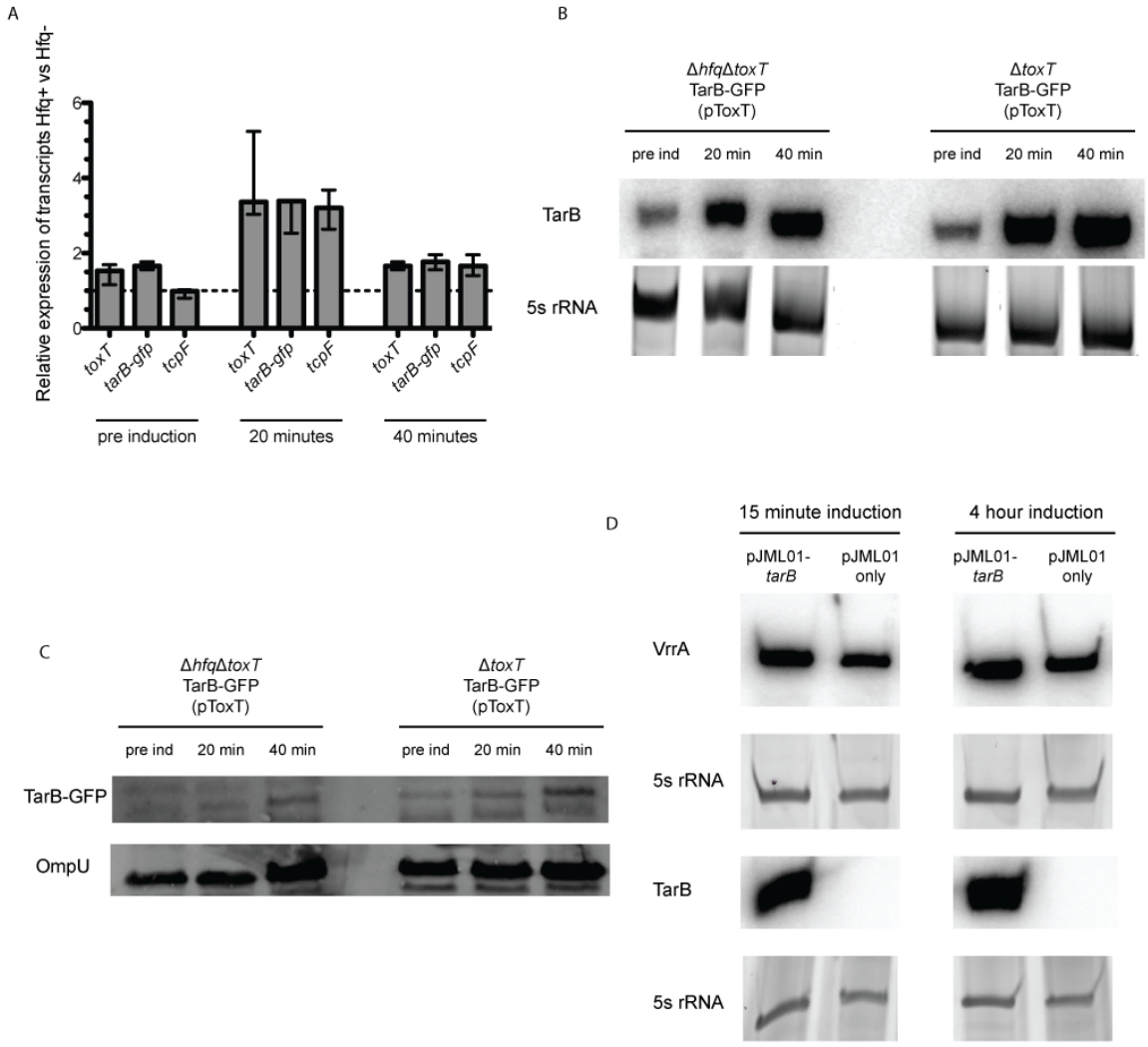

Figure 3.4: The RNA chaperone Hfq plays no detectable role in TarB stability or in its interaction with TcpF transcript.

A) The TarB promoter-GFP fusion was made in strains deleted for *toxT* and carrying arabinose-inducible ToxT on a plasmid in both the Hfq$^+$ and Hfq$^-$ backgrounds. These strains were then used to measure expression from the *tarB* promoter-*gfp* fusion, TarB from an intact native allele, expression of *toxT* from the plasmid, and expression of *tcpF*, which is a target of TarB by qRT-PCR. Data reported is the relative expression of those transcripts, adjusted for *rpoB* in the Hfq$^+$ strain relative to the Hfq$^-$ strain. Although

expression of all transcripts were higher at 20 minutes post induction in the Hfq^+ strain, the levels were similar before induction and after 40 minutes of ToxT induction. Adjusted for *toxT* expression, no differences were observed between Hfq^+ and Hfq^- strains for expression of *tcpF* and *gfp* from the *tarB* promoter-*gfp* fusion. **B)** A northern blot for TarB was carried out on the same RNA samples used in Panel A for qRT-PCR, the results indicate that there is no large difference in steady state level of the TarB sRNA in the Hfq^+ and Hfq^- strains, suggesting that Hfq has no role in stabilizing TarB. **C)** Results from Panel A were confirmed by western blot for GFP in samples taken from the same experiment. The results indicate that adjusted for loading, the two strains are expressing similar amounts of GFP from the *tarB-gfp* fusion prior to induction and at 40 minutes, indicating the *tarB-gfp* fusion is activated by expression of ToxT, as expected. **D)** To determine if over expression of TarB alters the stability of other sRNAs, we measured expression of the sRNA VrrA in cultures expressing TarB from the pJML01-*tarB* plasmid for either 15 minutes or 4 hours. In either case, expression of the VrrA sRNA was identical in cultures expressing TarB or those carrying vector only, indicating that over expression of TarB does not affect the stability of VrrA

TarB and TcpF expression during incubation in pond water

To determine if the expression of TarB and TcpF in the pond contributed to the reduced *in vivo* phenotype we observed in the Δ*tarB* mutant after pond water incubation, we carried out experiments to measure TarB and TcpF levels over the course of pond water incubation. The results of these experiments are summarized in Figure 3.5. TcpF expression was followed through the course of pond water incubation via the C-terminal FLAG fusion in both the wildtype and Δ*tarB* backgrounds by anti-FLAG western blot. The results indicate that the wildtype and Δ*tarB* mutant show similar levels of TcpF expression initially, however, over the course of pond incubation, TcpF levels drop in the wildtype strain, but not the Δ*tarB* strain. Transcription of TarB, as measured by production of GFP from the TarB promoter-GFP fusion indicates that levels of TarB expression do not change dramatically over the course of pond water incubation. The apparent upregulation of p*tarB*-GFP(ASV) at the 24 hour timepoint may relate more to the stability of the *gfp* mRNA and not increased transcription at that timepoint, as there is likely little metabolic activity occurring in these cultures after prolonged incubation in nutrient poor conditions. Northern blots for TarB expression over the course of pond water incubation suggest that TarB steady state level drops (Figure 3.5 panel C), but this may be due to the observed wholesale degradation of RNA after increasing time of incubation in pond water, such that accurate measurements of TarB expression via Northern blot may not possible. These results indicate that while TarB expression levels do not vary dramatically over the course of pond water incubation, TcpF protein levels do

drop, and this drop was absent in the Δ*tarB* mutant. This enhanced TcpF expression in the Δ*tarB* mutant may contribute to the phenotype of the Δ*tarB* mutant *in vivo* after pond water incubation, as over expression of TcpF in pond water would contribute to metabolic drain prior to infection or perhaps be responsible for some other subtle defect in fitness upon entry into the host.

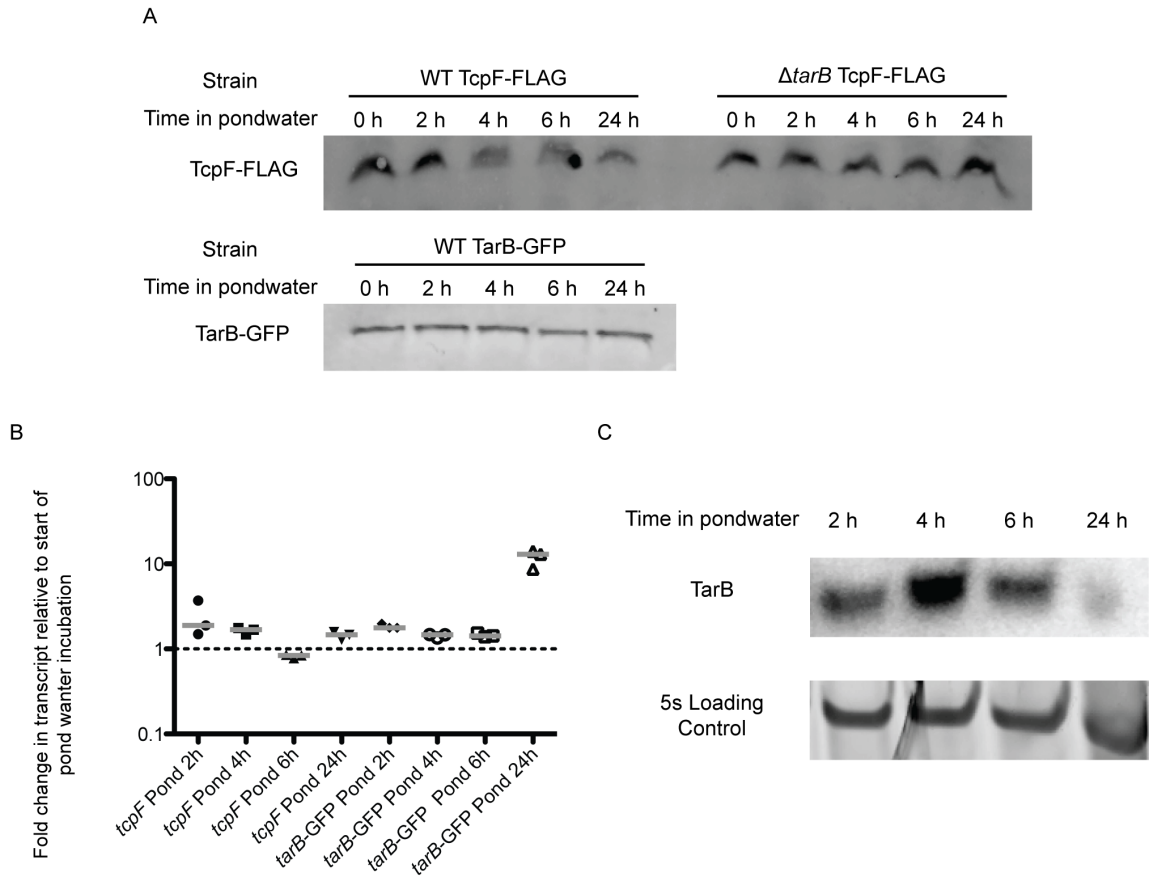

Figure 3.5 Expression of TcpF and TarB during pond water incubations.

A) Strains carrying the C-terminal TcpF-FLAG translational fusion or *tarB-gfp* (ASV) transcriptional fusion were incubated in pond water for the indicated amounts of time then lysed by boiling in SDS-loading buffer. Samples were then blotted with anti-FLAG and anti-GFP antibodies, loading was adjusted for OD. Levels of TcpF protein decline over the course of pond water incubation, this effect was absent in the Δ*tarB* mutant. B) Expression from the TarB promoter, as measured by GFP protein expression from the TarB-GFP fusion, however, does not vary greatly over the course of pond water incubation. C) Northern blot for TarB over the prolonged course of pond water incubation. Adjusted for loading, TarB does not show dramatic differences in abundance during the course of pond water incubation except at 24 hours when the signal is greatly

diminished. This, however, may be due to general degradation of RNA within the culture after prolonged pond water incubation.

In vivo expression of TarB

Although detectable fluorescence above background was produced by the TarB-GFP-ASV fusion during growth *in vitro*, bacteria with fluorescence above background were not detected by examination of tissue sections from infected mice (data not shown). To attempt to gain some insight into where and when TarB is expressed *in vivo,* infections were again carried out using a second reporter strain harboring a *tarB* promoter-*gfp* containing a point mutation in the RBS for *gfp* that eliminated detectable translation. This strain was a merodiploid such that the native TarB promoter and sRNA sequence was still present. Mice were sacrificed at various time points after infection, small intestines were split into proximal and distal halves, and RNA was extracted and used for qRT-PCR (Figure 3.6). In both the distal and proximal small bowel, TcpF and *tarB* promoter-*gfp* expression was measured. By using the expression level of these transcripts at 10 hours as the baseline and adjusting for loading of bacterial RNA by normalizing to *rpoB* transcript levels, we observed that expression of *tarB* and *tcpF* co-vary and both increase over the course of infection (see Figure 3.6). While this intuitively makes sense (the same system which activates TcpF expression also activates TarB expression**),** it does seem counter-productive that TcpF is upregulated at the same time that TarB represses it. This may indicate that TcpF expression at the level of translation is being tightly controlled by TarB. However, it is worth noting that this experiment yields information about the population average expressing *tcpF* and *tarB-GFP*, and thus variation at the level of the bacterial cell, if present, would be missed.

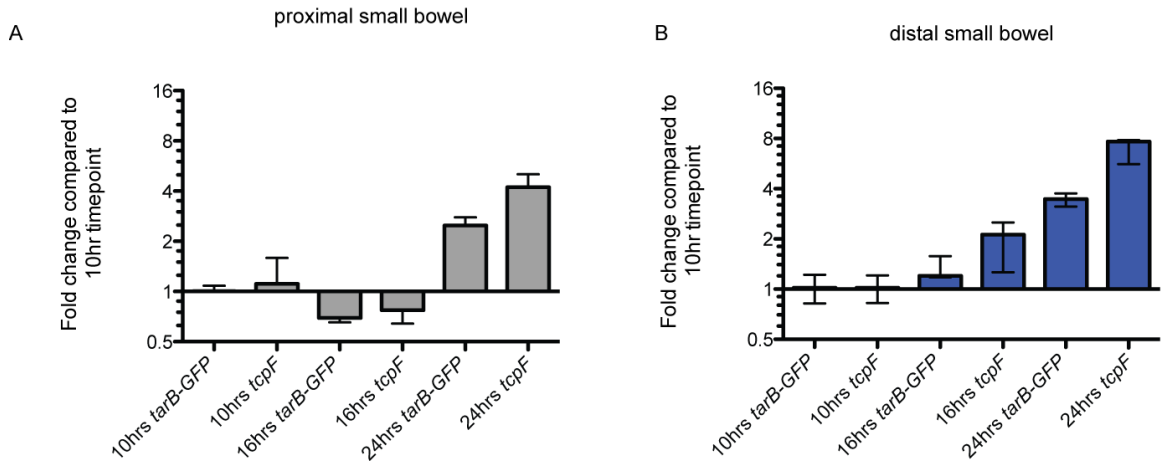

Figure 3.6 *In vivo* qRT-PCR of the *tcpF* and the *tarB-gfp* reporter

Infant mice were infected with a strain of *V. cholerae* harboring a transcriptional fusion of the *tarB* promoter to the *gfp* open reading frame. This construct contained a mutation in the ribosomal binding site, such that no detectable GFP protein was produced, but was instead used as a template for qRT-PCR to assess expression of TarB at various timepoints of infection. Expression of the *tcpF* message was also measured. At the indicated timepoints, mice were sacrificed and the small bowel was split into proximal and distal halves (**A** and **B**, respectively). Total RNA was extracted for use in qRT PCR. All reported values were adjusted for bacterial load and RNA input using *rpoB* as a housekeeping gene. Each reported value is the median of three mice.

Discussion

From the above work, it is clear that there are various inputs that modestly affect TarB expression, including anaerobic conditions, population density, and entry into stationary phase. The experiments we performed under anaerobic conditions suggest that oxygen plays a slight repressive role in TarB expression. TarB's function under low oxygen tension could be to repress TcpF expression prior to penetration of the mucous barrier of the small intestine. Upon reaching the epithelial surface, the higher oxygen tension would contribute to reduced TarB expression, allowing TcpF to be fully expressed. This would fit with the proposed role of TcpF in colonization of the epithelium [42]. The intestinal brush border is a highly vascular structure, commensurate with its role in absorbing nutrients, and it is reasonable to speculate that the luminal space adjacent to it would have greater oxygen tension then the luminal fluid. The actual oxygen tension of the small intestinal lumen may be quite low as oxygen-requiring luciferase reporter systems in bacteria do not function in the small intestine [172,173]. However, to the best of our knowledge, oxygen measurements at the brush border have not been reported.

The effect of population density on TarB via LuxO mediated repression may be relevant in the context of infection. *V. cholerae* is primed to express its virulence factors when the LuxO system is active at low cell density [78]. As we will describe later, TarB seems to be generally inhibitory towards the expression of virulence factors within the VPI-1. At the early stages of a *V. cholerae* infection, cell density would likely be low

and LuxO would be phosphorylated. Repression of TarB at this point may allow for early expression of virulence factors to allow for initial replication. As has been demonstrated numerous times, not all organisms during an infection show uniform virulence factor expression and there is temporal variation in when organisms are maximally expressing virulence factors [57,60], this level of TarB regulation may be important in generating this phenomenon.

During incubation in pond water, it does seem as though TarB is expressed at some level and that this serves some repressive effect against TcpF under these conditions, although this does not result in a survival defect of the Δ*tarB* mutant. This observation may however contribute to the observed virulence defect that the ΔtarB mutant has after pond water incubation. Perhaps inappropriate expression of TcpF during growth in pond makes the organisms less fit for a subsequent infection.

However, it is worth noting though that in the end, the most important factor influencing TarB expression in my experiments is ToxT. The differences seen in TarB expression under aneaerobic conditions or in the Δ*luxO* strain during growth in LB were on the order of 2-fold, whereas expression of ToxT from an arabinose inducible plasmid results in a 50-fold increase in expression. When the various *luxO* mutants were grown under virulence factor inducing conditions, it was clear again that ToxT was the primary driver of TarB expression, as any repressive effect that the *luxO** mutation may have had on TarB expression was over-ruled by increased expression of ToxT in that strain. Therefore, I propose that the additional environmental factors of oxygen tension, quorum sensing and entry into stationary phase, play relatively minor roles in regulating the expression of TarB, and that instead, ToxT plays the major regulatory role. Hence,

factors that influence ToxT expression and activity such as HapR mediated repression of TcpP/H will be the primary determinants of TarB expression.

Chapter 4 Investigating alternative targets of the TarB sRNA

Acknowledgements

All high-throughput sequencing was performed by the Tufts University Genomics Core Facility, all individual experiments were performed by Evan Bradley.

The existence of multiple diverse targets of regulation by individual sRNAs is an increasingly recognized and important paradigm. It suggests that the regulatory network of sRNAs is expansive and potentially highly complex. Several sRNAs in *V. cholerae* have been shown to have multiple targets [104,111] and we wondered if this was also true for TarB. We have previously shown that TarB negatively regulates TcpF primarily at the level of translation initiation by occluding the Shine-Dalgarno site and also modestly reduces *tcpF* mRNA abundance [92]. Additionally, it has been reported that TarB, dramatically affects mRNA abundance of *vspR* [12]. These observations open up the possibility that yet more factors may be directly or indirectly regulated by TarB. To test this, we used transcriptional profiling under TarB deletion and TarB over-expressing conditions during growth under *in vitro* virulence factor inducing condition [161] in an attempt to mimic the conditions under which TarB is usually expressed, to hopefully reveal additional targets that may be missed in LB growth. Here we show that TarB

indeed has additional targets predicted to be involved in virulence, nutrient uptake, and house keeping functions.

Determining alternative direct targets of TarB

To investigate direct and indirect targets of TarB within the host genome, we used strand-specific high throughput sequencing of cDNA [174] (RNA-seq) generated from cultures of *V. cholerae* grown under virulence factor inducing conditions [76]. We investigated the effects of artificial induction of TarB expression for a brief period (15 minutes) and over the course of the growth (4 hours). Cultures expressing an arabinose-inducible copy of TarB in a Δ*tarB* background were compared to vector alone at these two time points to determine what genes might be differentially regulated. This resulted in two data sets, genes immediately and thus possibly directly affected by TarB expression (Appendix Table 3, 15 min after induction), and those possibly indirectly regulated by TarB (Appendix Table 4, 4 h after induction).

Genes differentially regulated by at least two fold at 15 minutes or at 4 hours were then cross checked against potential interaction partners of TarB within the *V. cholerae* genome as predicted by the sRNA-target prediction program TargetRNA [153]. The same stringency of search parameters (see Materials and Methods) was used that yielded the interaction with TcpF, but potential interactions at 5' ends, 3' ends and within the coding sequences of possible mRNA targets were included. This resulted in nine predicted direct target mRNAs of TarB (including *tcpF* annotated VC0837). The annotation of these genes and their expression as measured by RNA-seq at 15 minutes and 4 hours is shown in Table 3. These genes were evaluated for expression after induction of TarB over the

course of growth in the virulence gene inducting condition AKI in three biological replicates using qRT-PCR (Figure 4.1, panels A and B). We included *tcpF* in this analysis to serve as a positive control. Data from individual genes correlates very well with the RNA-seq data, though interestingly, many genes show opposite changes in expression at 15 minutes and 4 hours.

Many of these potential target genes showed changes in expression at 15 minutes after TarB induction, suggesting that they are possible direct targets of TarB. The two predicted targets that showed the greatest changes in abundance after 15 min of TarB induction when followed up individually were VCA0686 (repressed 3.8 fold) and VC1863 (induced 4.5 fold). In the case of VC1863, this is particularly interesting as it was determined to be a gene regulated, possibly indirectly, by ToxT in a recent study of genome-wide expression in a ToxT over-expressing strain [12]. Thus, ToxT may act through TarB to enhance expression of this gene.

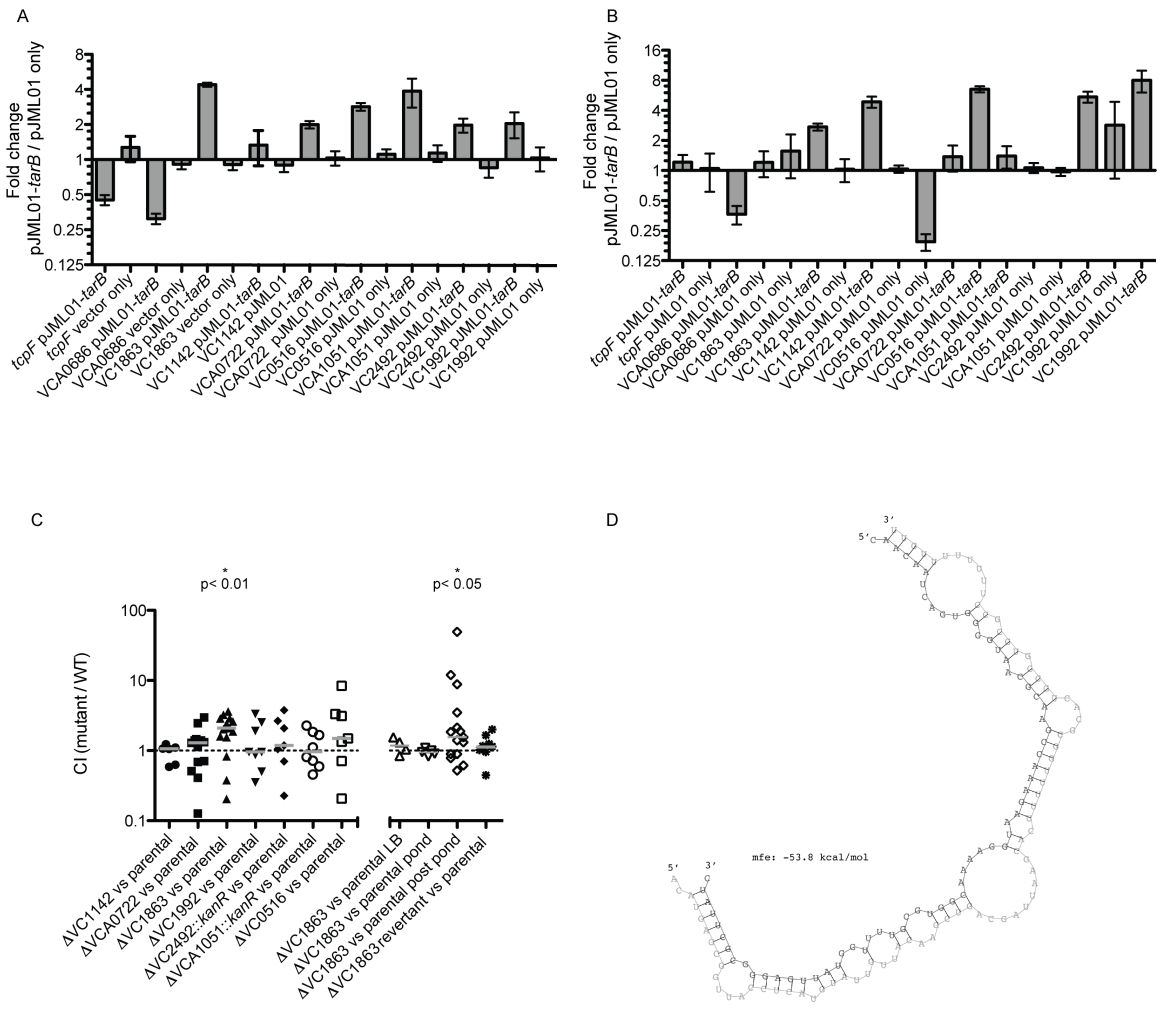

Figure 4.1 Determination of putative direct targets of TarB.

A) Measurements of mRNA steady state levels of putative direct targets of TarB by qRT-PCR after 15 minutes (panel A) and 4 hours (panel B) static growth in AKI at 37^{o}C and TarB induction. Measurements of the *tcpF* message (a known direct target of TarB) are included as a control. Shown is the mean and SEM of 3 biological replicates expressing TarB compared with vector alone. VC1863 and VCA0686 showed the most reliable changes in expression across experiments. C) The results of competition experiments carried out with strains deleted for genes that are putative direct targets of TarB. Only one

mutant (ΔVC1863) showed a significant deviation from a competitive index of 1 across experiments (Wilcoxon Signed Rank Test, $p < 0.01$). This increased colonization phenotype was not observed in a revertant of VC1863 D) The predicted duplex structure and mean free energy of pairing between TarB and the 5' coding region of VC1863 as predicted by the program RNAhybrid, the TarB sRNA is shown in darker gray.

Among other genes that showed reduced expression at 15 minutes of TarB induction were *tcpA* within the *tcp* operon, and *ctxA* and *ctxB* within the CtxΦ lysogen that encode CT. The *tcpA* mRNA and TcpA protein showed similar changes in abundance by qRT-PCR and Western blot, respectively, after 15 minutes of TarB induction at the endpoint of AKI growth (Figure 4.2). The cause of this repression does not appear to be reduced expression of ToxT in the TarB over-expressing strain, as both RNA-seq and individual analysis by qRT-PCR indicates that *toxT* is, if anything, more highly expressed (Figure 4.2 PanelA). Transcriptional regulators upstream of ToxT also show enhanced expression in the TarB over-expressing strain: *tcpP* and *tcpH* were significantly upregulated at both 15 minutes and 4 hours after induction, we confirmed this finding by qRT-PCR (Figure 4.2 panel D), and *toxR* and *toxS* were mildly upregulated at both times, though these increases were not statistically significant. These findings are consistent with higher levels of ToxT expression. This suggests that TarB exerts a repressive effect on the TCP operon independent of the major virulence gene regulators, however, no direct interaction was predicted between TarB and any gene of the pilus operon or within the VPI-1 other then *tcpF*. TarB must have a repressive effect on TcpA and the *tcp* operon independent of ToxT, this is the only possible explanation for decreased expression of these genes in the face of higher ToxT expression. A number of other virulence related genes were predicted to be upregulated by TarB, such as VCA0446 (the hemagglutinin gene) which we confirmed by qRT-PCR, again suggesting a complex series of regulatory events downstream of TarB

We investigated whether or not the binding site for TarB on the *tcpF* message could account for the repression seen of the entire operon. This was done by utilizing previously constructed *tcpF** and *tarB** strains harboring mutations within the 5' UTR of *tcpF* and TarB, respectively, that destroy the interaction between the *tcpF* mRNA and TarB [92]. However, no reproducible changes in abundance of TcpA protein or *tcpA* mRNA were seen in these strains (data not shown). This is consistent with recent transcriptional profiling data from infections of infant rabbits with *V. cholerae* showing that *tcpF* may have its own promoter [60,175] and thus may be part of a distinct mRNA. Hence the repressive effect of TarB on the TCP operon appears to be through some other mechanism.

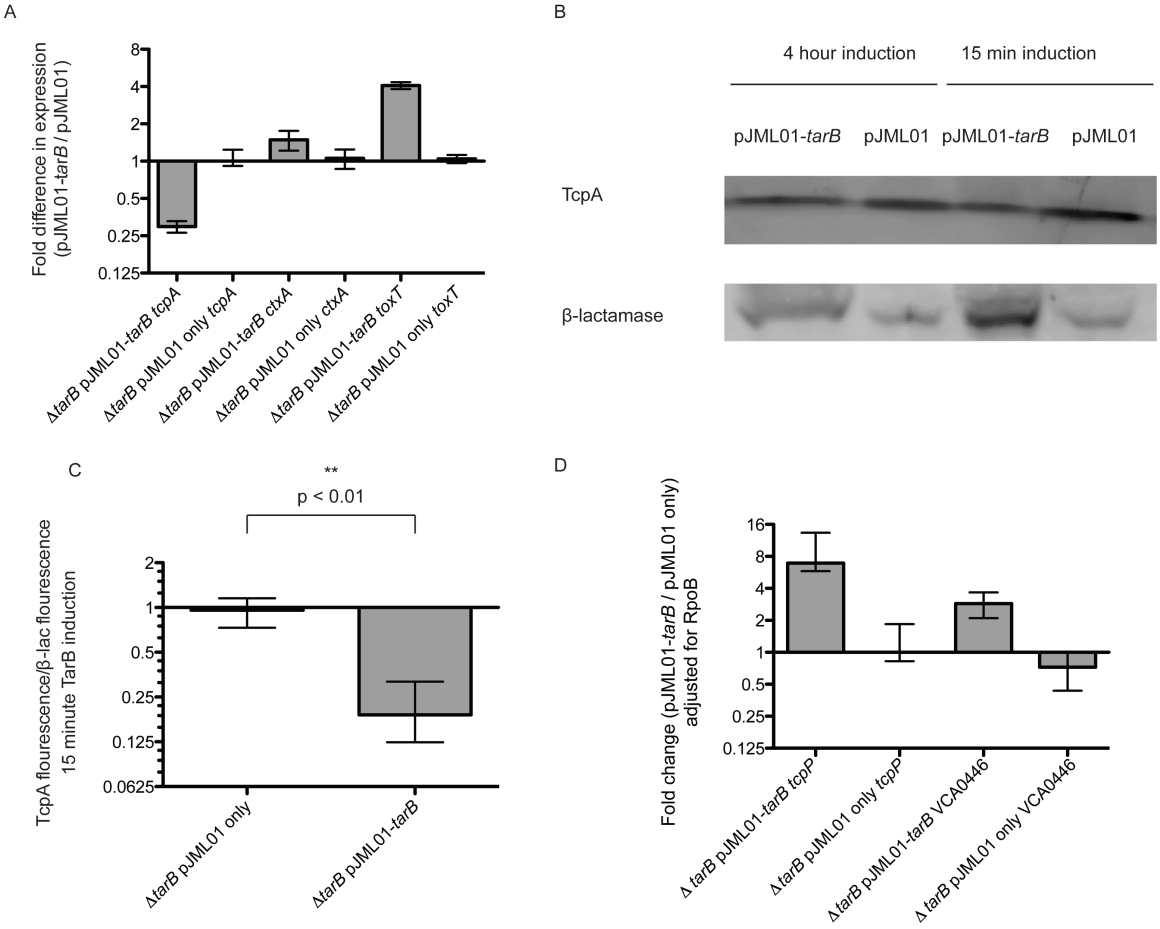

Figure 4.2 TcpA expression is inhibited by TarB expression.

A) Measurements of mRNA steady state levels of *tcpA*, *ctxA* and *toxT* by qRT-PCR after 4 hours static growth in AKI at 37^{o}C followed by TarB induction for 15 minutes. Despite *toxT* mRNA level being elevated in the TarB expressing strain, *tcpA* mRNA is down-regulated as was predicted by the RNA-seq experiment.The expression of *ctxA*, however, remains largely unchanged. B) Western blot for TcpA expression after the same culture conditions used in panel A. After blotting for TcpA, membranes were stripped and reprobed for β-lactamase expressed from the pJML01 plasmid to estimate total protein

loading. C) Quantitation of TcpA bands adjusted for β-lactamase loading control. Values were obtained by measuring total fluorescence of TcpA bands and dividing by the values for total fluorescence measurements of the β-lactamase bands in the corresponding lanes. Loading-adjusted TcpA fluorescence values were then normalized to measurements taken from the pJML01 vector only lanes. D) qRT-PCR for *tcpP* transcript and VCA0446 (hemagglutinin) transcript. After 4 hours of TarB induction, both are upregulated relative to vector alone as we would predict based on the RNA-seq data.

Several metabolic genes putatively regulated by TarB (Tables S1 and S2) were affected differently at 15 minutes than at 4 hours after TarB induction. Using analysis provided by the Biocyc website (biocyc.org) we determined the metabolic pathways that were upregulated in the TarB over-expressing strain. Many of these pathways appear to show different patterns of expression at 15 minutes than at 4 hours of TarB expression. For example, genes for maltose uptake and metabolism (*malE*, *malP*, *malQ*) and genes for glutamate biosynthesis are repressed immediately after TarB expression, but upregulated after 4 hours. How these time-dependent changes in genes affected by TarB expression occur remains to be investigated.

Pathways that were down-regulated at both 15 minutes and 4 hours of TarB expression include those involved in nucleotide metabolism and translation. It was previously shown using fluorescent protein fusions to the *rrnB* ribosomal RNA promoter that bacteria expressing the highest levels of *tcpA* also showed the highest level of expression from the *rrnB* promoter, which was inferred to mean this was the most rapidly growing population [60]. Since TarB appears to be repressive towards virulence gene expression, it may also repress genes involved in general growth, such as ribosomal proteins and pathways involved in amino acid biosynthesis. Arguing against this, though, was the observation that bacteria over-expressing TarB from the pJML01-*tarB* plasmid showed no defect in growth curve experiments when compared to vector alone (Figure 4.3). Because of these seeming conflicting results, the implications of the mild repression observed for nucleotide metabolism and translation genes after TarB over-expression remain unclear.

Although another group undertook a similar experiment to determine targets of the TarB sRNA, their findings do not entirely agree with ours [12]. Although one of their putative targets, VC2706, did appear as significantly regulated in our analysis, the other, VC0177 (also known as *vpsR*), did not. The reasons for this are unclear, different growth conditions (LB vs AKI growth) and our use of direct over-expression of TarB as opposed to measuring changes in a $TarB^+$ vs $TarB^-$ strain both over-expressing ToxT may explain some of the differences.

Roles of TarB targets in virulence

To determine if any of the genes identified as putative TarB regulated targets contributed to the previously observed increased colonization phenotype of the Δ*tarB* mutant, individual knockouts of several genes was constructed using natural competence and the Flp/Frt system [176]. Mutants in those genes were then used in competition experiments with the parental strain to determine if they contributed to the increased colonization phenotype of the Δ*tarB* mutant. Of the genes predicted to be direct targets of TarB, only one, VC1863, had a reproducible phenotype (Figure 4.1, panel C). VC1863 is part of a predicted amino acid ABC transporter. This deletion mutant had a mild, but statistically significant colonization advantage over the wild-type strain in competition experiments in the infant mouse model of colonization.

In a second experiment, in order to model a more natural infection model, the ΔVC1863 mutant was preincubated in pond water prior to infection of infant mice. The out-competition phenotype of this mutant was also seen after pond incubation (Figure 4.1, panel C). This mutant did not have an observable phenotype in *in vitro* competitions

or in a growth curve (Figure 4.1 panel C and Figure 4.3) suggesting that this increased colonization phenotype is specific to colonization of the small intestine. Our data indicate VC1863 is up-regulated by TarB expression. This phenotype is consistent with TarB acting as an anti-colonization factor. Some genes that are up-regulated by TarB expression would be predicted to be detrimental or inhibitory to colonization, and hence, deletion of these genes may provide a fitness advantage during infection, and that appears to be the case for VC1863.

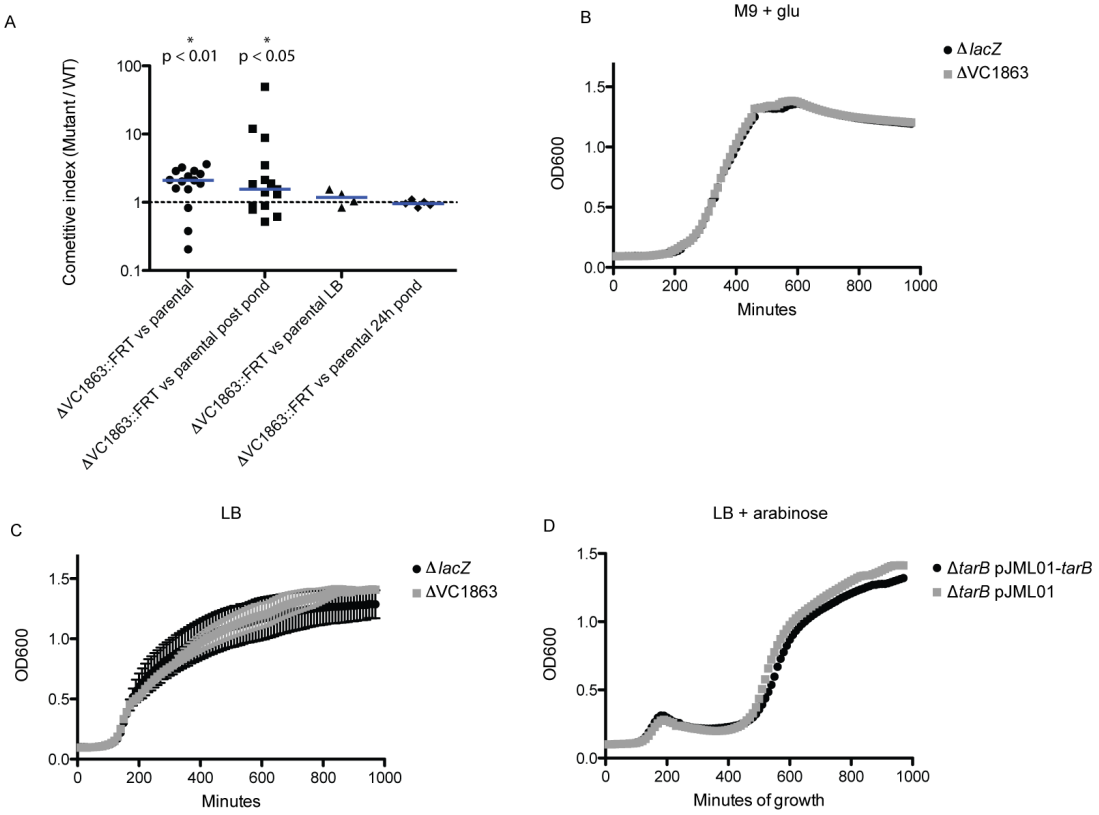

Figure 4.3 Further characterization of Δ1863 mutant and pJML01-*tarB* strain

A) Competitions with the ΔVC1863 mutant vs the parental strain *in vivo* after growth on LB plates and after 24 hour pond incubation, and competitions carried out *in vitro* in LB and in pond water. Only *in vivo* competitions after growth on LB plates and after 24 hours of incubation in pond are significantly different from a CI of 1 ($p < 0.05$, Wilcoxon-signed rank test). C/D) the ΔVC1863 strain shows no significant defect in growth curves in LB or minimal media when compared to the parental strain carrying a *lacZ* deletion. All data points are values reported are for biological triplicates. D) TarB expression from the pJML01-*tarB* plasmid does not appear to affect growth in LB as measured by growth curves.

Our lab has previously evaluated another predicted direct target of TarB, VCA0686 [21], which is predicted to encode the periplasmic component of an iron (III) uptake system. It showed no observable phenotype in *in vivo* competitions carried out previously in our lab. However, VCA0686 has been shown to contribute to survival of *V. cholerae* in a pond environment and be induced late in infection [58]. Given that TarB is up regulated by ToxT, a transcription factor that enhances expression of genes identified in our lab that are expressed early during infection, it seems logical that TarB may have a role in repressing genes important in the pond, but not necessary during the early phase of colonization and replication. In the case of VCA0686, TarB expression appears to down regulate this transcript. TarB is predicted to bind within the 5' UTR of this gene (Figure 3) and likely acts by a mechanism similar to how TarB binds and represses TcpF, by preventing translation at the start codon of the message.

In contrast, TarB appears to up-regulate expression of VC1863. The lowest energy predicted structure of the 5' protein coding region of the VC1863 mRNA as predicted by Mfold [177] reveals a region with one small and two large loops separated by short stems over which TarB has extensive complementarity (Figure 4.4). It is possible that, through binding of TarB to this 5' coding region of the transcript, expression of VC1863 is somehow enhanced, possibly by the unfolding of a potentially inhibitor structure shown in Figure 4.4.

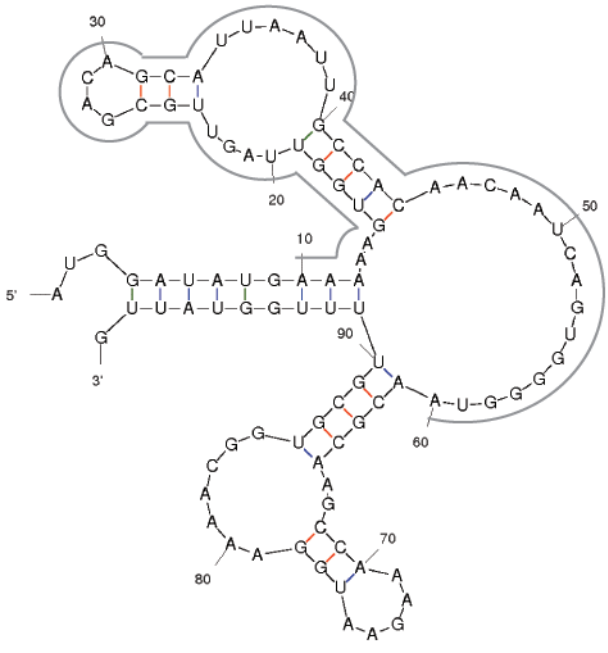

Figure 4.4 Mfold diagram of 5' region of VC1863 mRNA

Shown is the predicted folding of the 5' region of the VC1863 mRNA, the region TarB is predicted to bind is outlined in grey.

TarB directly interacts with the VC1863 and VCA0686 mRNAs

To determine if the changes in expression in VCA0686 and VC1863 are due to direct interaction of TarB with their mRNAs, we generated point mutations in each gene that are predicted to abrogate binding to TarB. Mutations in VCA0686 were designed so as not to alter the coding sequence or ribosome binding site (RBS) of that message. However, since TarB is predicted to bind with partial complementarity within the coding sequence of VC1863, some alteration to the amino acid sequence of the gene was unavoidable. We also generated compensatory mutations in TarB and cloned these mutant versions into the pJML01 vector. The mutations we constructed are shown in Figure 4.5, panels A and B. The level of expression of VC1863 and VCA0686 was then evaluated by qRT-PCR during growth in AKI media after 15 minutes of TarB induction (Figure 4.5, panels C and D). In the case of VC1863, when either TarB or the mRNA of VC1863 were mutated, abundance of that mRNA is reduced. However, when the compensatory mutations in TarB are present, the interaction is restored and the mRNA abundance is increased. In the case of VCA0686, because TarB appeared to repress that gene, abundance of the mRNA is increased when mutations are made to either the 5' UTR of the mRNA or the TarB sRNA, but again, when the complementary mutations are made in both the message and the sRNA, abundance drops, again suggesting that the interaction between TarB and this mRNA is direct.

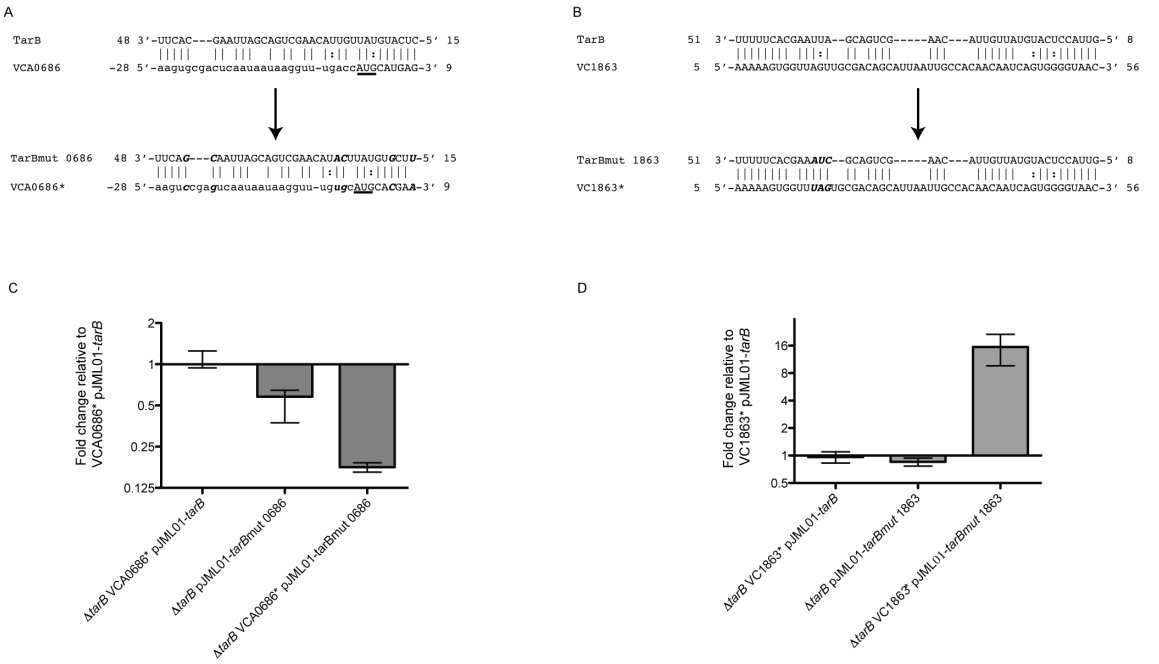

Figure 4.5 VCA0686 and VC1863 are direct targets of TarB

A and B) Shown are the interactions between TarB and the mRNAs of the genes VCA0686 and VC1863 as predicted by TargetRNA. Shown in bold are the mutations we made in either TarB born on the pJML01 plasmid or on the mRNA of the respective genes on the chromosome. C) D) qRT-PCR carried out on VCA0686 and VC1863, respectively, in the various mutant backgrounds after growth in AKI with induction of TarB for 15 minutes. In the case of VCA0686, repression is restored in the double, compensatory mutant strain. In the case of VC1863, mRNA levels increase in the double, compensatory mutant. This suggests that TarB affects these transcripts by direct RNA/RNA binding.

Discussion

Research into sRNAs in *V. cholerae* is increasingly showing that these sRNAs frequently have multiple targets [111,178]. TarB appears to be no exception, having at least four direct targets and numerous proposed indirect targets. These target genes are associated with a wide variety of cellular processes including virulence via regulation of TcpF and the *tcp* operon in general, chemotaxis as demonstrated by its effect on production of a recently described novel cyclic dinucleotide [12], and nutrient uptake and physiology as described here. The extent of our knowledge about the processes TarB regulates are shown in Figure 6. The cumulative result of these complex regulatory events appears to be an initial positive role in promoting colonization of the small intestine, but later a slight inhibitory role on net multiplication.

Previous work in our lab has demonstrated that many genes not directly associated with colonization and growth in the infant mouse model are expressed at this late stage of infection [58]. Many of these genes are hypothesized to be important in survival in the aquatic environment after release from the host. Two of the genes found to be significantly repressed by TarB were previously identified as “late genes” and were shown to plays roles in surviving the transition from the intestinal tract into pond water (ref 28). Integrating our knowledge about the targets of TarB and its colonization phenotype, we can begin to consider a model of TarB’s role during infection.

As has been previously established, repression of chemotaxis appears to be important to the infectivity of *V. cholerae* [62,65]. A recent study showed that TarB negatively regulates VC0177, and that VC0177 is a repressor of the dinucleotide cyclase

DncV [12]. DncV sythesizes the cyclic AMP-GMP dinucleotide, which represses the expression of chemotaxis genes [12]. Thus, through this chain of regulation, TarB appears to be functioning as a repressor of chemotaxis, which would be predicted to increase colonization of the small intestine. At the early stage of infection, it appears that TarB also represses expression of TCP and TcpF. Perhaps these colonization factors are not needed at the earliest stages of infection or must be tightly related via opposing action of the ToxT activator and the TarB repressor. In addition, as work here has shown, TarB regulates other genes that are not required for colonization but instead have been shown to be important for survival outside of the mouse (including VCA0686, VC1593, and possibly VC1863). Indeed, expression of at least one of these genes (VC1863) appears to be detrimental to colonization. All of these regulatory events therefore appear to be important for early colonization and replication.

However, as the infection progresses, it appears that the repressive effect of TarB on chemotaxis, *tcp* and/or other genes inhibits growth somewhat. Consistent with this hypothesis, recent work using a GFP fusion to the *tcpA* and *rrnB* promoters has suggested that those bacteria expressing TcpA to the highest level are also those bacteria located adjacent to epithelial cells lining the wall of the small intestine and that is the most rapidly growing *V. cholerae* sub-population [60]. Other studies have suggested that late in infection, chemotaxis genes are up regulated to promote detachment from the intestinal epithelium and entry into the lumen of the small intestine, or the so called “mucosal escape response” [49]. One potential hypothesis would be that at later time points of infection, TarB mediated repression of *tcp* and *tcpF* leads to inhibition of growth of these strains, and thus a Δ*tarB* population in competition experiment would expand more and

overcome its initial colonization defect as we have witnessed in the time course competition we have carried out here.

It is interesting to note that extended expression of TarB has different effects then a short course of expression of TarB. Repression of TcpA and the *tcp* operon was seen at 15 minutes, but not 4 hours, which was confirmed by Western blot. Thus, another possible explanation for the time-dependent virulence phenotype observed in the Δ*tarB* mutant is that extended expression of TarB later in infection leads to different regulatory events than when TarB is initially expressed early in infection. This could be due to positive feedback wherein decreased expression of genes mediated by TarB may lead to events that then cause those genes to be upregulated possibly over-riding TarB mediated regulation. The low numbers of bacteria present in the small intestine in the infant mouse model at early times of infection makes this hypothesis difficult to test using, for example, qRT-PCR on TarB regulated genes. Thus, we have not investigated this hypothesis experimentally.

Among sRNAs, TarB appears to be somewhat unusual. All of our investigations into the contribution of Hfq to TarB stability and activity have suggested that TarB acts independently of Hfq. The fact that TarB is Hfq independent has important implications for our RNAseq experiments, it suggests that changes in gene expression we see in our TarB over-expressing strain versus the null are not due to TarB occupying all free Hfq in the cell. However, the over expression approach may reveal lower affinity interactions between TarB and potential targets, these interactions may not be biologically relevant. Additionally, TarB's interaction VC1863 is somewhat unique, since it binds within the coding sequence of the mRNA, but seems to enhance its stability, rather then prevent its

translation or enhance its degradation it. Other sRNAs identified that bind within the coding sequences of messages are repressive [118], hence this may indicate a new mode of sRNA regulation.

Table 2 Putative targets of TarB determined by transcriptional profiling using RNA-seq.

Gene	Annotation	Log_2 fold change (pJML01-*tarB*/vector only) 15min	p value 15 min	Log_2 fold change (pJML01-*tarB*/vector only) 4 hours	p value 4 hours
VC0837	toxin co-regulated pilus biosynthesis protein F	-3.5	4.20E-14	-0.66	0.094
VC0561	Phage integrase	-0.23	0.64	1.27	0.0017
VC1992	formyltetrahydrofolate deformylase	-0.4	0.33	1.46	0.00039
VC1142	cold shock-like protein CspD	-0.27	0.53	2.38	1.04E-08
VC1863	amino acid ABC transporter, periplasmic amino acid-binding protein	-0.49	0.24	1.92	5.39E-07
VC2492	isopropylmalate isomerase large subunit	-0.48	0.24	1.24	0.024
VCA0686	iron(III) ABC transporter, permease protein	-1.75	0.0024	-3.1	1.78E-15
VCA0722	Hypothetical protein	-0.04	0.93	-1.48	0.0049

VCA1051	Hypothetical protein	-0.62	0.41	-1.54	0.00015

Table 3 Genes within the *Vibrio* pathogenicity island 1 operon affected by TarB expression.

Gene	Annotation	Log_2 fold change (pJML01-*tarB*/vector only) 15min	p value 15 min	Log_2 fold change (pJML01-*tarB*/vector only) 4 hours	P value 4 hours
VC0837	toxin co-regulated pilus biosynthesis protein F	-3.5	4.20E-14	-0.66	0.094
VC0828	toxin co-regulated pilin	-2.34	6.77E-08	0.26	0.51
VC0836	toxin co-regulated pilus biosynthesis protein E	-2.26	8.83E-08	-0.46	0.31
VC0834	toxin co-regulated pilus biosynthesis protein S	-2.19	9.40E-08	-0.2	0.65
VC0835	toxin co-regulated pilus biosynthesis protein T	-2.22	1.48E-07	-0.42	0.34
VC0827	toxin co-regulated pilus biosynthesis protein H	1.65	0.0046	2.3	1.48E-08
VC0826	toxin co-regulated pilus biosynthesis protein P	1.64	0.0056	2.2	3.33E-08
VC0824	tagD protein	-2.08	0.00028	0.32	0.43

Chapter 5 Perspectives and Future Directions

Future Directions

Here we have shown that new high throughput techniques can isolate sRNAs that were previously undetected and could not be predicted using bioinformatics or older sequencing methods [114,179]. In addition, we show the power of these tools to facilitate genome wide searches for genetic elements. Our experiments have expanded the list of possible sRNAs present in the *V. cholerae* genome and have expanded our appreciation of the diversity of possible sRNAs within bacterial genomes as being transcribed not only from intergenic regions, but from within ORF's and anti-sense to them as well. Still, we have only scratched the surface of transcriptional changes in *V. cholerae* that occur in the sRNA size range during ToxT expression. As was previously mentioned, direct sequencing and cloning in this experiment detected over twenty thousand potential sRNA transcripts. While it is certain many of these cloned products were the result of random RNA degradation, determining which transcripts are true sRNAs is an analysis we have not completed. That being said, degradation products of larger transcripts can still have biological activity, and some even require it [135]. This data set represents a large potential pool of sRNAs many of which not have been predicted by bioinformatic approaches, but the potential for false positives still exists. The data we have generated here could guide further studies into sRNA discovery as well as the contribution sRNAs make towards gene regulation during infection.

We have greatly expanded the list of putative ToxT binding sites within the genome and though there is some disagreement between our data and ChIP-seq

experiments performed by other researchers, it is clear that ToxT influences more then processes simply associated with expression of the classical virulence factors (the TCP, CT and the accessory colonization factors). Either directly or via downstream regulators, such as TarA and TarB, ToxT definitely has an impact on the physiology of the organism. As was mentioned before, determining which of these binding sites are relevant *in vivo* and which make significant contribution to gene regulation *in vivo* remains to be determined.

Prior to this work, the existence of sRNAs under the control of the ToxR regulon were hypothesized [114], but had never been isolated or observed. We now understand, through our work and the work of two other labs [144], that not only do such sRNAs exist, but one, TarB regulates the production of a novel cyclic di-nucleotide [12] and has important effects on virulence gene regulation during infection [92]. Further study of TarB and its downstream targets could yield valuable insights into spatial and temporal expression patterns of virulence genes and help us understand better how the infectious process is orchestrated by *V. cholerae* to the detriment of the host. Despite the work here, a definitive explanation of TarB's function *in vivo* remains elusive, though we do now have an in-depth analysis of what it regulates. While we have elucidated much about the regulation and targets of the TarB sRNA, its dynamic expression during infection remains to be investigated. We generated a number of different transcriptional fusions of fluorescent proteins to the TarB promoter, but none yielded detectable fluorescence in sections of small intestine taken from mice infected with these strains, though they did colonize to approximately wild-type levels (data not shown). Given the time-dependent phenotype of the Δ*tarB* mutant and its regulation of genes that appear to be important for

survival outside of the host environment, understanding where and when TarB is expressed would be critical to understanding its role in infection and to help us better understand the infectious cycle of *V. cholerae*.

Transcriptional profiling of a TarB over-expressing strain compared to a null yielded a large number of genes that were differentially regulated between the two conditions. While direct binding could be predicted and validated for two of these targets, the mechanisms by which the other genes are regulated, including the *tcp* operon, remains unknown. Clearly TarB is effecting more then just virulence gene expression, but what regulators are acting down stream of TarB? Are these targets transcription factors or are they affecting mRNA abundance by some other mechanism? There is clearly more to be done to understand the factors downstream of TarB.

Among sRNAs, TarB appears to be unusual. Its stability and activity appear to be Hfq independent, which is not commonly seen among *trans*-acting sRNAs. It also appears to regulate one of its targets (VC1863) by binding within the coding sequence of that gene, another unusual feature. These findings highlight the diversity of sRNA-mediated regulation in bacteria and suggest that new mechanisms and modes of regulation by sRNAs remain to be discovered. It is possible that TarB acts through as-yet undetermined mechanisms to regulate mRNA and protein levels, since it does not utilize the classic Hfq machinery that many other sRNAs act through. Does it utilize other chaperones instead? Or is this an unusual case of a sRNA acting completely on its own? Investigating the mechanisms of TarB mediated regulation may reveal new insights into how sRNAs regulate gene expression in bacteria.

Because of these unusual findings with respect to how TarB works, it maybe beneficial to determine what proteins TarB interacts with, if any. The use of RNA aptamer tags would be particularly useful in this kind of study [180]. An aptamer tagged TarB allele could also be used as biochemical evidence of TarB's interaction with the binding partners we elucidate in this study, strengthening the genetic evidence we have already put forward.

Biochemical evidence of TarB's interaction with the targets determined here would be beneficial beyond simply confirming previous results. The genetic technique of mutating a sRNA of interest, then mutating its target, and restoring the interaction between the sRNA and the target by making the two mutations in the same bacteria is powerful evidence of an interaction between the two molecules, but it is not without potential fault. For example, mutating the sequence of the sRNA or the target could disrupt secondary structure critical to function outside of the base paring interaction that allows the two to interact, and may result in off target effects which may affect the phenotype under study or make a negative result difficult to interpret. Co-purification of target mRNA with aptamer tagged sRNAs or the observation of duplex formation by alerted RNA mobility on native gels [130] would provide additional evidence for interaction without potentially altering the sRNA or mRNA function.

Many unexplored avenues exist to determine what contribution TarB makes and how other sRNA's control gene expression during infection in *V. cholerae*. Though we and other labs have begun to determine what genes are controlled by sRNAs in V. *cholerae*, the data we have generated seems to suggest that there is a layer of control and complexity to expression of virulence factors that has been previously unappreciated.

sRNAs as members of the ToxR regulon

In the course of this work we have uncovered two new sRNA members of the ToxR regulon in *V. cholerae*, called TarA and TarB. Both are expressed from genes within the *Vibrio* Pathogenicity Island-1 (VPI-1), an island that also harbors genes for the TCP and virulence gene regulators. TarA and TarB were discovered independently by other groups using different methods [12,144]. These two sRNAs take part in the regulation of important processes during infection in animal models and they highlight some important features of sRNA-mediated regulation.

A key facet of regulation by sRNAs is that it can occur very quickly. In an examination of the genes controlled by ToxT by RNA-seq by another group, it was shown that VC1863 was upregulated after 15 minutes of ToxT expression [12]; our work here suggests that this upregulation of VC1863 occurs via TarB. This would indicate that steady state levels of VC1863 could be increased without altering transcription of the gene after only 15 minutes of expression of the primary factor upstream of TarB (ToxT). Although direct targets of ToxT were also upregulated at this time, TarB may have the advantage of acting on existing mRNAs. Within 2 minutes of induction of the TarA sRNA, expression of its target, *ptsG* mRNA is reduced [144]. Such quick responses may provide a fitness advantage during the highly dynamic and stressful infection process.

Another interesting feature of TarB is it's ability to regulate a very recently acquired genetic element. The *Vibrio* Seventh Pandemic island (VSP-1) is a gene cluster that appeared to have emerged in the *V. cholerae* genome within the last century [10], and yet TarB, which is encoded in the more ancestral VPI-1, is capable of regulating a

gene within this island. Recent publications suggest that sRNAs are emerging as important regulators of genes acquired by horizontal transfer relatively recently in evolutionary time. The mutation of a single base-pair within one of two homologous genes acquired by horizontal transfer in *Salmonella typhimerium* resulted in discrimination between the two transcripts by the SgrS sRNA [181]

This exemplifies what may be a unique ability of sRNAs to rapidly acquire new targets. In order for a protein-based transcription factor to acquire a new target gene, the amino acid sequence of the protein would need to mutate, or alternatively a DNA element within the promoter of the target gene would need to mutate to allow an existing regulator to bind. Such mutations in a hypothetical protein regulator would likely be a very low frequency event, as it would require not only the proper amino acid change(s) to recognize a new promoter, but the mutations must not adversely affect the folding or stability of the protein, nor affect its binding to pre-existing targets. On the other hand, mutation within the promoter region of the new target gene that results in the recruitment of a pre-existing regulator is likely to be more relatively more frequent. But what about sRNAs acquiring new target mRNAs? Mutations occurring within the 5' UTR of an mRNA that allow for targeting by a pre-existing sRNA should be on par, in terms of frequency, with the mutations in promoter sequences that would allow binding of new regulators. Both would require that mutations occur within the appropriate location within the promoter/5' UTR. However, perhaps the most frequent of all would be the occurrence of mutations within pre-existing sRNAs that direct it to new target mRNAs. Such an event was hypothesized to occur in the *hld* gene within the *Staphylococcus aureus agr* locus [182]. This locus is thought to have mediated biofilm formation in non-

pathogenic strains of *S. aureus*, however the *hld* RNA transcript acquired mutations that allowed it to become an RNA regulator of a large number of virulence-related genes and it became a central regulator of virulence in some strains virulent strains of *S. aureus.* This hypothesis is based on the nature of targeting by sRNAs, which typically occurs through base pairing to mRNAs over relatively short sequences of pseudo-complementary. Thus, new mutations could occur in in positions within the sRNA that lie outside (or even in non-critical positions within) the pseudo-complementary regions in order to gain a new target, but which do not adversely affect binding to the pre-existing target. This may be how TarB evolved to control a recently acquired gene cluster.

TarA and TarB also highlight co-evolution between mobile genetic elements and the core genome of pathogens. The VPI-1 is hypothesized to, at least at one time, have been a mobile element, though experiments to prove this have been controversial [43,183]. Despite this, TarA and TarB, which are encoded within the VPI-1, appear to regulate genes outside the VPI-1 and within the core genome of *V. cholerae*, again highlighting the ability of sRNAs to readily acquire new targets. TarA and TarB join a growing list of sRNAs encoded by mobile elements in various bacteria that affect expression of core genes [140,141,142].

Understanding TarB's role during infection

The work we present here paints an interesting if sometimes contradictory picture of what the role of TarB is during infection. While TarB appears to be generally inhibitory towards virulence factor expression (e.g. repressing TcpA and TcpF), and to

chemotaxis as well [12], this regulation appears to be important only under a particular set of conditions, those being early in infection and when coming from a resource poor environment. However, as the infection progresses, TarB's repression of these genes, seemingly counterintuitively, results in reduced net replication of *V. cholerae* within the host. But, perhaps there is a logical explanation to this finding. Previous work in our lab using the infant mouse model has characterized 24 hours as a "late" timepoint during infection, in which many genes that do not affect *V. cholerae's* ability to replicate in the host begin to be expressed [58]. Some of these genes are instead involved in dissemination to the aquatic environment, a step which is clearly important to the life cycle of this pathogen. Thus, perhaps TarB is needed to repress virulence genes at this late stage of infection in order to allow a subset of bacteria to switch from a "colonization/replication" mode of gene expression, to a "dissemination/nonreplicating" mode of gene expression, and that is why we see the phenotypes we do.

A diagram of the current state of our knowledge based on our experiments and other work on TarB is shown in Figure 5.1. It appears that quite a few processes are affected by expression of TarB and the impact of TarB expression on those metabolic pathways remains to be investigated, however, it does lend some insights into what pathways may be important in the host. For example, maltose utilization genes have previously been shown to be activated by ToxT [12]. They are however, repressed in the TarB expressing strain at 15 minutes (see Appendix Table 3) and activated after 4 hours of TarB expression (see Appendix Table 4). This may suggest that TarB initially downregulates maltose utilization early in infection, but that this regulation is reversed later in infection, hence different carbon sources may be utilized at different times.

A process that is consistently affected by TarB expression is translation. A large number of ribosomal proteins are downregulated in the TarB expressing strain at both 15 minutes and 4 hours of expression. This may indicate that TarB is repressing this process, perhaps to allow for transition into the hypothesized dissemination state. Another process that may regulate growth rate is TarB's affect on amino acid metabolism, which is affected differently depending on how long TarB is expressed. Although expression of TarB showed no effect on growth *in vitro* (Figure 4.2), the contribution of TarB expression to growth rate *in vivo*, where growth conditions are likely to be subobtimal, remains to be investigated.

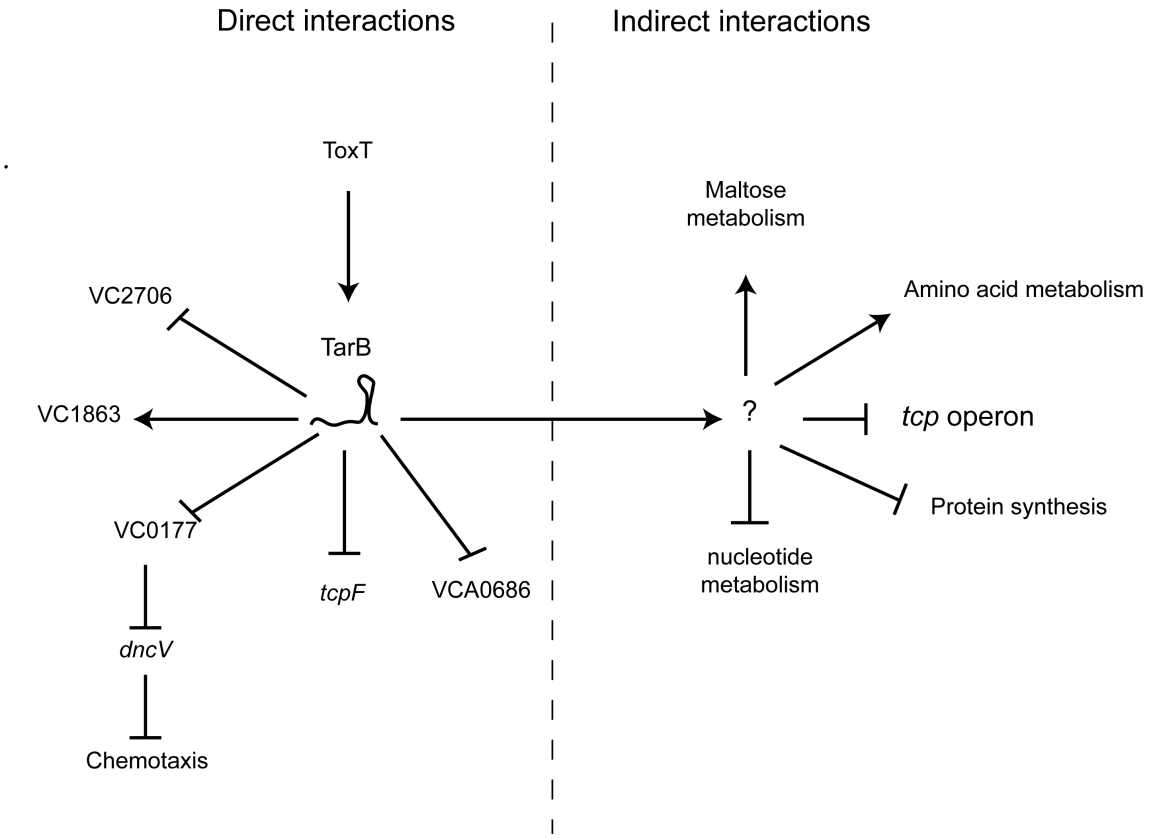

Figure 5.1 A map of TarB's down stream regulated genes

Shown on the left are the genes that TarB has either been shown to directly regulate here or elsewhere. Positive interactions are denoted with an arrow, negative interactions are denoted with a blocked line. Downstream genes include the putative amino acid ABC transporter VC1863, the putative iron (III) uptake system VCA0686, *tcpF,* potentially VC2706, and VC0177, also known as *vspR* , which is a regulator of genes within the VSP-1. Additionally, shown on the right are the results our transcriptional profile experiment, which suggest that TarB, through one or more regulatory intermediates, affects expression of the entire *tcp* operon, as well as effecting metabolic genes involved in maltose metabolism, amino acid metabolism and nucleotide metabolism. These results

suggest a far more complex regulatory network influenced by TarB then was previously appreciated.

Chapter 6 Materials and Methods

Bacterial growth conditions

V. cholerae O1 serogroup El Tor biotype isolate E7946 and derivatives were grown at 37°C in LB broth with aeration. For AKI induction, strains were grown in AKI broth (1.5% peptone, 0.4% yeast extract, 0.5% NaCl, 0.3% $NaCHO_3$) statically for 4 h at 37°C followed by aeration for 4 h 37°C. To induce expression of cloned genes on plasmids, arabinose was added to 0.04% (for pToxT derivatives) or 0.1% (for pJML01 derivatives) upon reaching mid-exponential phase (optical density at 600 nm [OD] = 0.3). All DNA manipulations were done in *E. coli* DH5α or derivatives with plasmids maintained with the appropriate antibiotics.

Strain construction

All PCR reactions were carried out with EasyA polymerase according to the manufacturer's specifications using the indicated primers, the sequences of which can be found in Table 4. The descriptions of all plasmids used in this study are included in Table 6.

Plasmids pToxT and pToxT ΔHLH plasmids we constructed by PCR

amplification of the *toxT* ORF including native RBS from gDNA from either wildtype *V. cholera* E6749 or an E6749 strain carrying an internal deletion of the helix-loop-helix DNA binding domain [57] using primers NcoI_ToxT_F and XbaI_ToxT_R. This PCR product was then cloned into the NcoI and XbaI sites of the pBAD24 plasmid [184] to allow expression of ToxT upon addition of L-arabinose. Unmarked deletions of chromosomal genes were constructed by SOE PCR introduced using a derivative of the pCVD442 allelic exchange vector, pCVD442-lac which contains the pUC19 LacZ gene and MCS, as described [185].[185].

Point mutations in the *tarB* gene were generated by SOE PCR using primers xbaI_TarB comp_F, TarB_mut_R1 and TarB_mut_F2 and SacI_TarB_comp_R, using E6749 genomic DNA as template. PCR products were mixed in a one to one ratio, and added to a PCR reaction run for 25 cycles at an annealing temperature of 50^{o}C without primers and the mutated sRNA sequence plus promoter were amplified with XbaI_TarB_comp_F and SacI_TarB_comp_R which contain SacI and XbaI restriction sites which were subsequently used for cloning into pMMB67EH to generate p*tarB**. The wildtype complementation vector p*tarB* was generated by cloning a PCR product generated using XbaI_TarB_comp_F and SacI_TarB_comp_R primers and genomic DNA as a template.

Point mutations in the *tcpF* 5' UTR were also generated by SOE PCR using primers XbaI_TcpF_mut_F1, TcpF_mut_R1, TcpF_mut_R2 and XbaI_TcpF_mut_R2 using an identical procedure as above. The final ~2kb product containing the mutated *tcpF* 5' UTR sequence which was subsequently cloned into the XbaI site of the pCVD442-lac vector which was then mated into strains of interest. Double crossovers

were selected on 10% sucrose plates. Individual double crossovers were screened for the mutated sequences by sequencing with the TcpF seq primer and the XbaI_TarB_comp_F primer and confirming double crossover by streaking on 10% sucrose as well as ampicillin containing plates to ensure sucrose resistance and ampicillin sensitivity. C-terminal FLAG fusions to TcpF were generated by amplification of the C-terminal 346 bp using the TcpF_qt_F primer and the TcpF-FLAG_R primer to add the FLAG amino acid sequence [186], this product was subcloned into Topo pCR2.1 (Invitrogen). The resulting plasmid was cut using KpnI and EcoRV and the insert containing the C terminus of TcpF with the FLAG fusion was cloned into a modified pGP704 suicide vector [187] which contains a chloramphenicol resistance drug marker in place of an ampicillin marker (pGP704cat). This construct was then mated into strains of interest and single crossovers were selected for on chloramphenicol plates at 2 μg/mL. Proper insertions were confirmed by PCR using the TcpF-FLAG reverse primer and TcpF seq forward primer. A merodiploid strain was constructed by plasmid integration resulting in the placement of GFP(ASV) under the control of one copy of the TarB promoter followed by the native TarB locus downstream of the integrated plasmid sequence. The plasmid borne fusion was generated by amplifying the +3 to -376 positions in the TarB promoter from E6749 genomic DNA using primers TarB_F and TarB_-300_R and subcloning the product into pCR2.1 yielding p*tarB-300.* GFP was amplified from pGfpmut3.1 plasmid (Clonetech) using primers Fgfp2 and Rgfp2 which adds a ribosomal binding site and SacI site at the 5' end and the destabilizing (ASV) [158]C terminal amino acids and a SmaI site at the 3' end. The GFP(ASV) PCR product was cloned in a triple ligation with the SacI/EcoRV fragment from p*tarB-300* into pGP704cat digested

with SmaI to generate the transcriptional fusion. The resulting plasmid (pTarB-GFP) was mated into E6749 strains and single crossovers were selected on chloramphenicol and confirmed by PCR using primers Rgfp2 and XbaI_ΔTarB_R2.

sRNA deep sequencing

Single colonies of strain AC3763 (Δ*toxT*) transformed with either pToxT or pToxTΔHLH plasmids were picked and grown in LB broth containing streptomycin and ampicillin both at 100 μg/mL overnight. Strains were back diluted from overnight cultures to an OD of 0.03 in 200 mL LB supplemented with streptomycin and ampicillin both at 100 μg/mL and were grown with aeration at 37°C until the strains reached mid-exponential phase (OD = 0.3). Arabinose was then added to 0.04% to induce expression of *toxT* alleles from pToxT plasmids, and induction was allowed to proceed for 20 minutes prior to RNA extraction. Total RNA was purified from the bulk culture by phenol/chloroform extraction and isopropanol precipitation. Cloning and sequencing of sRNA was carried out as previously described [114], sequences of the micro RNA cloning linkers (IDT) used are included in table S4. In order to further decrease tRNA and 5S rRNA in the final sequenced products, the depletion step described in the previously published procedure was carried out twice with the addition of an oligo targeting the serGCC tRNA (5'-GCGGTGAGTGAGAGATTCGAACTCTC-3'). The final cDNA products were prepared for Illumina Genome Analyzer II sequencing using Illumina primers 1a, 1b and 1c (table S1) for the first 10 cycles of PCR, followed by gel purification and Illumina primers 2a and 2b (table S1) for the final 4 cycles of PCR followed by PCR clean up (Stratagene). Final products were run on a Bioanalyzer High-Sensitivity DNA chip (Agilent) prior to

Illumina sequencing to normalize loading of the two samples and ensure quality of the libraries. The libraries were pooled and placed on one lane of an Illumina Genome Analyzer IIx paired-end sequencing run at Tufts University Core Facility. Briefly, a paired-end sequencing run sequences both the 5' and 3' end of every DNA molecule attached to the flowcell. The first read is downstream of linker 1 and the second read is downstream of linker 2 (ToxT library) or linker 3 (ΔHLH library) so that for every pair, the directionality of the original RNA molecule could be determined. Sequence reads were trimmed to remove linker sequences and filtered so that 100% of the sequenced bases in each read had a minimum quality score of 5 (base call accuracy at least 68%). Reads were aligned to the O1 biovar N16961 genome (NCBI Accession Nos. NC_002505, NC_002506) using Bowtie (http://bowtie-bio.sourceforge.net). Reads matching rRNA or tRNA regions were removed from the alignment, leaving 1,062,048 reads in the ToxTΔHLH library and 2,212,216 reads in the ToxT library. Unique transcripts totaled 6,815 for ToxTΔHLH and 27,787 for ToxT. The alignments were then processed to generate a library of clustered transcripts using the method previously described [114]. This resulted in 3,309 clusters for the ToxTΔHLH library and 12,534 clusters for ToxT library. Clustered reads were output into "gff" format and viewed using GenomeView (http://genomeview.org). The number of reads in sRNA clusters were normalized by dividing the number of reads in that cluster by the ratio of MtlS reads in that library to total MtlS reads. For example normalized $reads_{ToxT}$ = cluster $reads_{ToxT}$/ ($MtlS_{ToxT}$/($MtlS_{ToxT}$ + $MtlS_{ToxTΔHLH}$)). In the final output table, a cutoff was made of at least 500 normalized reads between either library. This generated a list of 765 potential sRNA transcripts that represent the most abundant transcripts in either library which may

represent true sRNA transcripts.

ToxT overexpression and purification

E. coli strain BL21(DE3) was transformed with the plasmid pMAL-TEV-His-thr-ToxT (table s3). The resulting strain was grown on LB agar plates containing ampicillin and a single colony was picked for growth of a 4 mL overnight culture. The overnight culture was used to inoculate 1 L LB broth containing ampicillin at 100 μg/mL and was grown with aeration at 37°C. Transcription was induced once the culture had reached exponential phase (OD=0.5-1) by addition of IPTG to 1 mM. Induction was allowed to proceed shaking at 20°C for 16 h, after which, cell pellets were collected by centrifugation and resuspended in 20 mL lysis buffer (20 mM Tris-HCl pH8, 2 mM DTT, 1 mM EDTA, 250 mM NaCl) plus Complete protease inhibitors (Roche). Cell pellets were lysed and the lysate was cleared by centrifugation at 18,000 rpm in a SS34 rotor.

The cleared lysate was then applied to a 5 mL dextrin MBPtrap column (GE Life sciences). The column was washed with lysis buffer followed by elution with MBP elution buffer (as lysis buffer, + 1 mM maltose). The elution fractions were subsequently diluted 10-fold with buffer QB1A (20 mM Tris-HCl pH8.0, 1 mM DTT) and applied to an 8 mL Source15Q anion exchange column equilibrated in QB1A. The protein was eluted using a 0 to 20% gradient of QB1B (20 mM Tris-HCl pH8.0, 1 M NaCl, 1 mM DTT) developed over 25 column volumes. The peak fractions were diluted 5-fold in SB1A buffer (25 mM phosphate buffer pH6.0, 1 mM DTT) and applied to a 8 mL Source15S cation exchange column equilibrated in SB1A. The protein was eluted using a 15 to 35% gradient QB1B (25 mM phosphate buffer pH6.0, 1 mM DTT, 1M NaCl),

which resulted in two peaks, the second peak was known to be a soluble aggregate and was discarded. The initial peak was split into two aliquots, one of which was applied to a Superose 12 gel filtration column in EMSA buffer (10 mM Tris-HCl pH7.5, 200 mM KCl, 10 mM βME) for use in mobility shift assays, the other aliquot was cleaved with TEV protease overnight at 4°C and subsequently diluted 5-fold in SB1A and applied to a 2 mL Source15S cation exchange column to separate His-ToxT from the cleaved MBP fusion protein. His-ToxT was eluted from this column with a 35 to 100% gradient of SB1B developed over 12 column volumes. Finally, His-ToxT peak fractions were applied to a Superdex 75 gel filtration column in EMSA buffer. These final steps did leave a small amount of TEV protease in the final purified product.

Affinity purification of ToxT binding sequences

Genomic libraries were prepared by centrifuging 10 mL of overnight growth of wild type (AC53) *V. cholerae*, washing 2x with TBS and resuspending in 5 mL TBS. To generate gDNA fragment sizes of 300 to 1,000 bp, the cell pellet was subjected to four 30 second sonication cycles on ice using a sonicator micro tip (Branson); each sonication cycle was separated by a 30 second incubation on ice. After sonication, RNAase A was added to a concentration of 2 μg/mL, the samples were incubated at 37°C for 20 min to allow for degradation of RNA. DNA was purified with 2 rounds of extraction with citrate buffered phenol:chloroform (Ambion) followed by a final extraction with chloroform only and then concentrated by ethanol precipitation. Fragmented DNA was used to prepare three different bar-coded libraries using adapters BC1a/BC1b, BC2a/BC2b and BC3a/BC3b (Table 5) as described [151]. For the final amplification and purification of bar-coded

libraries, ten PCR reactions were done using linkered and size selected gDNA as template using primers Olj 139 and 140 and EasyA polymerase (Stratagene). PCR conditions were as follows, denaturation for 5 minutes at 95°C, annealing for 30 seconds at 65°C, elongation for 30 seconds at 72°C, cycling back to denaturation at 95°C for 30 seconds for 15 cycles after which reactions were pooled and incubated with 50 μL ExoSAP-IT (USB) at 37°C for 1 h. Final purification of libraries was carried out by phenol:chloroform extraction and ethanol precipitation and resuspension of libraries in 100 μL deionized water.

Binding reactions contained 15 μg bar-coded DNA library in a total volume of 250 μL with 200 nM purified His6-tagged ToxT purified as above or with His6-tagged TEV protease in EMSA buffer with 10 μg/mL sheared salmon sperm DNA, 0.3 mg/mL BSA and 10% glycerol. Reactions were allowed to incubate with gentle mixing at 37°C for 1 h, after which the reaction was added to a microcentrifuge spin column (Pierce) packed with a 50 μL bead volume column of HisPur cobalt resin (Pierce) that had been equilibrated in the above buffer. The reaction was allowed to bind to the column by mixing gently at 37°C for 1 h. Flow through was then collected by spinning the column in a microcentrifuge at 3,000 x g for 1 minute. The column was washed 3x by gentle resuspension of the bead volume in 250 μL of EMSA buffer with the above additions, followed by centrifugation. The column was washed an additional 3x as above, but in EMSA buffer only. After the final wash, the bead volume was resuspended in 10 mM Tris-HCl pH8 and boiled for 5 minutes and allowed to cool to room temperature, then incubated with proteinase K (5 μg/ml) for 30 minutes at 65°C, followed by boiling for 5 minutes. After centrifugation for 1 min at 3,000 x g, the resulting 100 μl of the

supernatant fluid was purified by using a PCR purification kit (Qiagen) and then subjected to 10 cycles of PCR amplification with primers Olj139 and Olj140, repurified, quantified on the Bioanalyzer high sensitivity DNA chip (Agilent), and subjected to deep sequencing, along with aliquots of the input libraries prior to pulldown, using the Illumina Genome Analyzer II on the paired end setting.

Reads from the Illumina libraries were aligned to the N16961 genome. Sequence alignment and assembly were performed as described above. After alignment, reads that did not match the genome were discarded and the sets were normalized so that each set contained the same number of reads. Alignment positions were shifted by half their insert length as determined by each mapped pair, giving the center position of each sequenced DNA molecule. These positions were then tabulated and used to generate a coverage map of the genome using a rolling average with window size of 35 bases. Coverage maps were generated for every sample. For each genomic DNA and corresponding pulldown sample, an enrichment map was created, which represented the ratio of the values from the pulldown sample over that of the genomic DNA sample. Enrichment maps were then scanned to identify regions that had more than 3x the average coverage for more than 100 consecutive positions. The false discovery rate (FDR) was then calculated by performing the same analysis with the control and pulldown samples switched. At 3x coverage, the FDR was 0.03 and 199 enriched sites were identified totaled between the libraries, of which 67 were observed in both replicates. Significance of each enriched region was assessed using two methods [188]. First, the number of reads in that region in the control sample was used to generate a Poisson distribution. This was then used to assess the probability of the same number of reads occurring in the pulldown sample. Using this

method, all regions identified had a p-value of < $1x10^{-98}$. Second, a Z-score was found by comparing the proportion of tags in the control sample to that in the pulldown. All of the regions identified had a significant difference in the proportion of tags counted between the control and pulldown samples, with z-scores > 7.7. The nucleotide sequences from the overlapping set were used as a training set for finding motifs using MEME 4.1.0. We allowed MEME to find motifs that occurred at least one time in each fragment. The motif reported in Figure 2.1 Panel C is the lowest E-value motif for the 67 sites combined in both libraries.

Mobility Shift Assays

Primers TarB promoter R and TarB promoter F were used to amplify the upstream 100 bp of TarB, predicted to contain promoter elements and ToxT binding sites to serve as a probe in the mobility shift assay. The PCR product was purified (Stratagene) and 3.3 pmoles was end-labeled using T4 Polynucleotide Kinase (NEB) and ^{32}P γ-ATP according to the manufactures instructions, and then purified using a Performa DTR spin cartridge (Edge Biosciences). A negative control probe of similar size consisting of 4.5S RNA sequence was prepared in parallel. The binding reaction occurred in 20 μL with 3 nM labeled probe and varying concentrations of purified MBP-his-thr-ToxT in EMSA with 10 mM 10 μg/mL sheared salmon sperm DNA, 6 μg/mL BSA, 10% glycerol and 0.002% Orange G dye added. Binding was allowed to occur for 30 minutes at 30°C followed by loading of the entire reaction onto a 5% TBE-Polyacrylamide gel, which was then run at

100V for 60 minutes. The gel was then used to directly expose a phosphor screen and the image was read on a FLA-9000IR using the IP setting.

AKI induction experiments

For AKI induction experiments, strains were grown overnight with aeration at 37°C in LB broth containing streptomycin at 100 μg/mL, and the appropriate antibiotics for vector maintence (ampicillin at 50 μg/mL for pMMB-based plasmids, ampicillin at 100μg/mL for pBR based palsmisd). For strains carrying the TcpF C-FLAG integration in the wild type background and TarB-GFP strains, chloramphenicol at 2 μg/mL was included as well. Overnight cultures were then diluted into prewarmed AKI media [161] containing 0.3% $NaHCO_3$ and ampicillin at 50 μg/mL (again excluded for the wild type background strain and TarB-GFP fusions) to an OD of 0.01. Strains were grown statically in an incubator at 37C for the indicated times at which culture aliquots were removed for analysis. After 4 hours of static growth, cultures were split into 1 mL aliquots and grown shaking at 37C for 4 hours. For induction of pJML01-based plasmids during AKI growth, L-arabinose was added to a final concentration of 0.1% either at the end point of static growth, or it was included in the medium during growth

Anaerobic growth

For anaerobic growth experiments, overnight cultures were prepared by inoculation of strains into phosphate buffered LB media containing 60 mM K_2HPO_4, 33 mM KH_2PO_4, 0.5% glucose and 100 μg/mL streptomycin. These cultures were grown overnight in an anaerobic chamber and used to subsequently inoculate either 2 mL aerated cultures or 10

mL cultures in sealed tubes prepared in the anaerobic chamber to an OD of approximately 0.01. Aerobic and anaerobic cultures were then grown in parallel in a shaking 37C incubator to approximately the same OD and snap frozen on liquid nitrogen and subsequently used for RNA extraction and northern blots. For each culture the pH of the media was measured after growth was recorded and ranged between 6.3 and 6.5 for anaerobic cultures and 6.7 to 6.8 for aerobically grown cultures.

Pond Water Incubations

For pond water incubation experiments, strains were grown overnight on M9 minimal media + glucose plates containing the proper antibiotics. Overnight growth was resuspended in saline and washed twice. After the final wash, strains were resuspended in filter-sterilized pond water and inoculated into 2 mL culture tubes of filter sterilized pond water to an OD of 0.1 and incubated shaking at 37^{o}C for the indicated times. At those times, culture aliquots were prepared either for western blot by centrifugation followed by resuspension in sample buffer and boiling or diluted to a density of 1 x 10^{3}/μL as measured by OD for mouse infections.

ToxT induction experiments

Experiments involving induction of ToxT from the arabinose inducible plasmid were carried out similarly to those used in sRNA sequencing experiments. Overnight cultures of the indicated strains were grown at 37°C overnight in LB containing the appropriate antibiotics. Overnight cultures were then diluted to an OD of 0.03 in 25 mL of the same media and allowed to grow shaking at 37°C. Once cultures reached mid-exponential

phase (OD =0.3), arabinose was added to a final concentration of 0.04% and induction was allowed to proceed for 1 h with 2 mL aliquots of culture taken at the indicated times and either spun down for western blot analysis or snap frozen in liquid nitrogen for RNA extraction later.

Northern Blots

Between 2.5-10 μg of total RNA purified using the Ambion mirVana kit from the indicated cultures was run on 10% TBE-urea polyacrylamide gels. Prior to transfer, gels were stained with GelStar (Invitrogen) and scanned on the FLA-9000IR (Fuji) to assess total RNA loading in each well and to use for normalization during quantification. RNA was transferred to Hybond N+ membranes (Amersham) in 1x TBE using the Mini Trans-Blot Cell apparatus (Bio-Rad) according to the manufacturer's instructions. Blots were prehybridized in Ultrahyb (Ambion) prior to addition of probe. RNA probes were transcribed from PCR-derived templates with T7 promoters using ^{32}P-UTP and T7 polymerase (Promega) according to the manufacturer's instructions. Ambion Decade ladder labeled with ^{32}P-ATP was run alongside RNA samples to provide estimations for the sizes of RNA bands. Hybridzation was carried out at 65°C overnight followed by washing 3x with low stringency buffer (2x SSC + 0.05% SDS) wash at room temp, followed by washing 3x with high stringency buffer (0.2xSSC + 0.05% SDS) at 65°C. Blots were then exposed to phosphor storage screens (Fuji) overnight. The image was subsequently read on a FLA-9000IR scanner. When reporting quantification, measurements taken from the phosphor screen after exposure were divided by fluorescent measurements of the 5S rRNA taken prior to transfer to normalize signal for loading

using the MultiGage software (Fuji).

qRT-PCR

Total RNA was purified from cultures grown under the indicated conditions using the mirVana RNA purification kit. Total RNA was treated with DNAase with the TURBO-DNAfree kit (Ambion) prior to reverse transcription. cDNA used as template was generated using iScript complete kit (BioRad) from 2 μg of total RNA using random hexamers. Quantitative PCR was run using Strategene Mv3005P equipment and MxPro qPCR software. Each sample was measured in technical triplicate. In all cases, controls lacking reverse transcriptase were included to assess DNA contamination, all results were either below the baseline of detection, or were subtracted from values obtained with those templates.

Western Blots

For western blot analysis of TcpF and GFP expression, strains carrying the TcpF C-terminal FLAG allele or the TarB-GFP fusion were grown under the indicated conditions at which times 2mL culture aliquots were removed. Culture aliquots were immediately centrifuged at 10,000 x g for 5 minutes to pellet cells, and supernatants were removed. Cell pellets were boiled in 50 μL (static timepoints and plasmid induction experiments) or 100 μL (4 h aeration timepoint) of SDS loading buffer (50 mM Tris-HCl, pH6.8, 2% SDS, 0.5% bromophenol blue, 10% glycerol, 100 mM βME). Samples were cooled and a

volume adjusted for differences in OD was loaded on an SDS-polyacrylamide gel electrophoresis (PAGE) gel and run 90 minutes at 125 V. Proteins were transferred to a nitrocellulose membrane at 25V for 1 h. Membranes were loaded onto the SNAP-ID Western blotting system (Millipore) and blocked with 1x NAP blocking agent (G Biosciences) diluted in PBS + 0.01% Tween-20. Primary antibody to the FLAG peptide (Invitrogen) or against GFP (Abcam) was added to the membrane 1:600 or 1:1200 respectively, diluted in 3 mL 1x NAP block for 10 minutes and the membrane was washed with 90 mL PBS + 0.01% Tween-20. Secondary antibody (Invitrogen) (Cy5 conjugated goat anti mouse for anti-FLAG blots or Cy5 conjugated goat anti rabbit for GFP blots,) was added to the membrane at 1:600 and diluted in 3 mL 1x NAP block for 10 minutes and the membrane was washed with 90 mL PBS + 0.01% Tween-20. Bands were visualized using the Cy5 setting the FLA-9000IR. After visualization of TcpF-FLAG, blots were stripped by incubating in 20 mL acid stripping buffer (25 mM glycine pH2, 1% SDS) shaking for 30 minutes followed by washing 2x with 20 mL PBS + 0.01% Tween-20. After stripping, blots were reprobed as above with primary anti-OmpU at 1:600 in 1x NAP block and secondary Cy5 conjugated goat anti-rabbit (Invitrogen) again in 1x NAP block and scanned on Cy5 setting on the FLA-9000IR.

Fluorescence measurements were quantified using MultiGage software (Fuji). Measurements TcpF-FLAG bands, adjusted for area and background, were divided by fluorescence measurements of corresponding OmpU bands adjusted for area and background. Loading-adjusted fluorescence values were then standardized to wild type expression and reported as fold expression of TcpF relative to wild type expression. The experiment shown is representative of six biological replicates.

Western blots to determine TcpA expression in strains harboring pJML01 based plasmids during AKI growth were performed as above with the following exceptins. The primary antibody (rabbit anti-TcpA) was at a 1/3000 dilution in NAP block made in TBS/T (Tris-buffered sailine with 0.01% tween)

Mouse infections

Single strain infections and competition assays in infant mice, LB broth and filter sterilized pond water were performed with the TarB unmarked deletion strain (AC3744) ($LacZ^+$) and wild type with a *lacZ* deletion (AC3745) for 24 h as described [189].[189]. Inputs for competition assays and single strain infections were prepared by growth overnight on LB plates containing the appropriate antibiotics followed by resuspension in LB to an approximate density of 1 x 10^3/μL as measured by OD, mixing of equal volumes of either culture (for competition experiments) then inoculation of infant mice by oral gavage. Samples from pond water incubations were prepared as described above, mixed in equal volumes and then used for innocualtion of infant mice. Immediately after inoculation, input ratios and total CFU were determined by plating on LB plates containing 5-bromo-4-chloro-3-indolyl-D-galactopyranoside (X-gal). The target input dose for all experiments was 10^5 bacteria/mouse, although over the course of the experiments doses ranged between 10^4 and 10^6. Results are shown by the competition index (CI), which is the ratio of mutant CFU to wild type CFU normalized for the input ratio. To show complementation *in trans* in all assays in this study, Δ*tarB* derivatives ($LacZ^+$) were complemented with either p*tarB* or p*tarB** and were competed against the

respective isogenic strain (LacZ$^-$) carrying the pMMB67EH plasmid alone. CIs for these experiments are expressed as the ratio of mutant to complemented CFU corrected for input. To assess plasmid loss frequency, output plates were replica plated onto LB agar plates containing streptomycin and ampicillin at 100 μg/mL and X-Gal at 40 μg/mL to determine plasmid containing CFUs, and LB agar plates containing streptomycin and X-Gal to determine total CFUs.

Growth curves

Growth of strains was determined by measured OD using a Bio-Tek microplate reader. Cultures grown overnight in LB plus streptomycin and (ampicillin at 50μg/mL for complemented strains) or M9 glucose plus streptomycin and (ampicillin at 50μg/mL for complemented strains) were resuspended to an OD of 0.01 in the respective media and pipetted into a 96-well plate in triplicate. Each growth curve was performed in biological triplicate. Bacteria were grown with aeration for 17 h at 37°C in the microplate reader with the OD being read every 17 minutes.

RNA-seq experiments

For each RNAseq experiment, three biological replicates were prepared from cultures carrying pJML01-*tarB* or pJML01 only after induction for the entire course of AKI growth or for 15minutes at the end point of AKI growth. Strand specific cDNA libraries were prepared for Illumina sequencing using the dUTP labeling method as

described by Levin et al [174] with some modifications.Strand specific cDNA libraries were prepared for Illumina sequencing using the dUTP labeling method as described by Levin et al [174] with some modifications. Total RNA was extracted using the RNAeasy kit (Qiagen) and was treated with Microbe Express (Ambion) to deplete ribosomal RNA. RNA was then sheared by sonication in a Branson Sonifier cup sonicator (ask Ayman about this) at maximum amplitude for 2 minutes total with 10 second pulses on and 5 seconds off.

Sheared RNA was then reverse transcribed using SuperScript III RT (Invitrogen) using a modified protocol. The reaction was carried out with 250ng of random hexamers with a 1 hour extension time at 55C and no heat denaturation step. RT reactions were cleaned up on Performa spin columns (Edge Biosciences). Second strand synthesis was carried out with the given units of the following enzymes, 2U RNAaseH, 40U *E. coli* DNA polymerase I, 10U *E. coli* DNA ligase in NEB 10x Second Strand Synthesis buffer supplemented 4mM DTT and 800 μM dNTPs with dUTP substituted for dTTP for 2.5 hours at 16C. Second strand synthesis reactions were then cleaned up with a PCR clean up column (Stratagene). Double stranded cDNA was then blunted with NEB quick blunting kit, A tails were then added using Exo – Klenow fragment (NEB) and Illumina TruSeq adapters were ligated with the NEB quick ligation kit. Final linked cDNA products were then gel purified or treated with Aline Size Selector beads to isolate fragmetns between 200 and 400bp in length. Size selected, linked cDNA wa then treated with 1U of USER enzyme (NEB) to remove dUTP residues incorporated into the second strand. USER was heat inactivated and cDNA was amplified using primers Olj344 and Olj335 and Phusion DNA polymerase in High Fidelity buffer (NEB). Libraries were

sequenced on the Illumina Hi-Seq at the Tufs University Genomics Core Facility

Analysis of transcript abundance was performed to compare cultures expressing TarB from vector alone using the CuffDiff [190] software package to perform RPKM analysis. Each set of 3 biological replicates was compared to the corresponding replicates carrying vector only, a cut off of $p < .01$ was used to determine genes that were significantly differentially regulated between the strains.

Table 4 Primers used in this study

Name	Sequence
ToxT_NcoI_F	GCCCATGGTATCTTCAGAGTAGAACGCAATGA
ToxT_XbaI_R	GCTCTAGATTATTTTTCTGCAACTCCTGTCA
IL1a	CTCTTTCCCTACACGACGCTCTTCCGATCTGATTGATGGTGCCTACAG
IL1b	GGCATTCCTGCTGAACCGCTCTTCCGATCTGTCCTTGGTGCCCGAGTG
IL1c	GGCATTCCTGCTGAACCGCTCTTCCGATCTGCTGGAATTCGCGGTTAAA
IL2a	AATGATACGGCGACCACCGAGATCTACACTCTTTCCCTACACGACGCTCTTCCGATCT
IL2b	CAAGCAGAAGACGGCATACGAGATCGGTCTCGGCATTCCTGCTGAACCGCTCTTCCGATCT
ΔtoxT_F1_(XbaI)	GCTCTAGATTCTCTGCTCGGCTTTTAGC
ΔtoxT_R1	GTAAACGTATTCCATTACATTGCGTTCTACTCTGAAGATATATA
ΔtoxT_F2	CAGAGTAGAACGCAATGTAATGGAATACGTTTACTTGATCCTA
ΔtoxT_R2_(XbaI)	GCTCTAGATTTGACACATCGACCTTGGA

ΔtoxT_F0	CTACGGATTCAAGGGGGAG
ΔtoxT_R0	TCCTGAACGTCATCTAGTGGT
toxT_NdeI_F	GCCATATGATTGGGAAAAAATCTTTTCAAACTAA
toxT_BamHI_R	GCGGATCCTTATTTTTCTGCAACTCCTGTCAACAT
MtlS_Rev	CCGTTGGTGATTCCATTCG
MtlS_For	TCCCCCGTTGGATGTTCCG
ΔtarA_F1-SphI	GCGCATGCTCGCTTGTATGTTTGGACGA
ΔtarAR1	TTCGGTTTAGCACTCCCTAACTTTATTTTCCTAAAGACAAA
ΔtarA_F2	GGGAGTGCTAAACCGAATGAATTATAATGAGAATTACTTT
ΔtarA_R2-SphI	GCGCATGCCCCCCAAGCTTTTAATTTTT
T7_tarB_F	TAATACGACTCACTATAGGGCGCCAAAAAGTGCTTAATCG
tarB_R	AAAACAAAATCATCTTTCATAACAGC
ΔtarB_F1_XbaI	GCTCTAGAGTGTTGGTGCTGCACACTCT
ΔtarB_R1	AGCAATGTAACCAACCTCAAATATTAACCCTTAGGATATTC
ΔtarB_F2	TTGAGGTTGGTTACATTGCTTTTTAACGCTCTTGTTTCTATTTAAGC
ΔtarB_R2_XbaI	GCTCTAGACCTTTCCCAAATTGAGTTCG
Δhfq_F1_xbaI	GCTCTAGAACTGATTTATCGAGGGATGG
Δhfq_R1	GATCCAGAAATGGGTCTTGTAGAGATTGCC
Δhfq_F2	CCCATTTCTGGATCGTCCAGCAGAGAAGTCT
Δhfq_R2_xbaI	GCTCTAGATTACGCAAAGTAGGATCGAG
T7_tarA_F	TAATACGACTCACTATAGGG CCAAACGTAAGGGGCAAAAT
tarA_R	ATAATTCATTCGGTTTAGCACTCC
tarA_comp_F	GATGTGAAAAATCAGCTTTTATCGT
tarA_comp_R_	ATTTGCAATCTAATTCTGCAGTTG
xbaI_tarB_ comp_F	GCTCTAGATTGAGGTTGGTTACATTGCTATAA
sacI_tarB _comp_R	GCGAGCTCGCTTAAATAGAAACAAGAGCGTTAAAA
TarB_promot	TGTATGTTTATAGTGCCAGTAT

er_F	
TarB_promoter_R_-100	CATAAGCTTAAATAGAAACAAGA
TarB_promoter_R_-300	TTTAAAGATAGAGTGATCGCG
4.5s_F	CTGGTCCTCCCGCAACAC
4.5s_R	GAGACCCCAGCCACATC
TcpF_mut_F1_-_xbaI	GCTCTAGAGAGGGAGTGGGCATCTATGA
tcpF_mut_F2	
	TTCTAGTTTATAGTGAGGTATTATGAGATATAAAAAAACCTTAATG
tcpF_mut_R1	ATACCTCACTATAAACTAGAACTTAGTTTATCAACGAGCG
tcpF_mut_R2_-xbaI	GCTCTAGACCGTTAAGTTGCCACTAGGC
tcpF_mut_F0	TGAAAATTATCTCCAAGAAGTATAGGC
tcpF_mut_R0	TTGACCACTTGTAACCATTATGC
tarB_mut_R1	CATGATATGTTACAAGCTGACCTATAAGCACTTTTTGGCGCACTGCGG
tarB_mut_F2	TTATAGGTCAGCTTGTAACATATCATGAGGTAACCGCTCATGTATG
tcpF_qt_F	TGGTGCAATGATCGCAGTAT
tcpF_qt_R	CCGTTAAGTTGCCACTAGGC
Vca_0638_qt_F	CGGTTTAGTGCGCCATTATT
Vca_0638_qt_R	CCATACACTTCCGCCAGAAT
Vc0177_qt_F	TAACGGTGAAGGGAGTGGTC
Vc0177_qt_R	TGGTTCCAGTTCAGGGAATC
Vc0937_qt_F	TTGGTTGATGTGCAAGGTGT
Vc0937_qt_R	TCAGCGACTTTCAAATCACG
Vc2506_qt_F	CAGCCAAGCTCAACAAAACA
Vc2506_qt_R	CATCAAACAGGCTCAAAGCA
CadC_qt_F	TATGTGGTGACGGTGCCTAA
CadC_qt_R	TTCGGCTTGCTTGATTTCTT
tcpF_C-FLAG-_R	TTAGCTCTAGTTATTTGTCATCGTCATCCTTGTAGTCTTTAAAGTTCTCTGAATATGC
Olj139	AATGATACGGCGACCACCGAGATCTACACTCTTTCCCTACAC

	GA
Olj140	CAAGCAGAAGACGGCATACGAGATCGGTCTCGGCATTCCTGCTGAAC
Fgfp2	GCGAGCTCTTTAGGATTTATTAAAATGCGTAAAGGAGAAGAACTT
RGFP2	GCGCCCGGGCTAAACTGATGCAGCGTAGTTTTCGTCGTTTGCTGCAGGCCTTTTGTATAGTTCATCCATGCC
nheI *tarB* F	GCGCTAGC ACATGAGCGGTTACCTCATG
sphI *tarB* R	GCGCATGC AACAAAAAAAAGGCGCACCGC
Olj 335	CAAGCAGAAGACGGCATACGAGAT
Olj 344	AATGATACGGCGACCACCGAGATC
Vca0686 qt F	CCTACTCTAGCTTCTGTATGCTGG
Vca0686 qt R	GAATGATGCTGATCAGTGAGCC
VC1142 qRT F	GTTTTATTTGCCCGGAAGGT
VC1142 qRT R	TTGCTGATTGTCCCTCGATT
VCA0722 qRT F	GATCCAACAGGCGAGCTTAC
VCA0722 qRT R	GTTTGTTCCAATTCGGCTGT
VCA1051 qRT F	GACATTGGTACCAGCGGTTT
VCA1051 qRT R	TGGCACCACAGTTTTACCAA
VC1863 qRT-F	GAGATGCAAGTGGAGTGCAA
VC1863 qRT-R	ACGCCCTTCAAACCTTCTTT
VC2492 qRT-F	TGGCGATCATTGGTAAAACA
VC2492 qRT-R	TTGCCAATAGTCGACAGCAG
VC1992 qRT-F	CCACTGTGTGTCCCATGAAG
VC1992 qRT-R	GGCAAGAAACTGTGGTGGAT
VC0516 qRT-R	GCGAATATGGCGCTAAAGAG

VC0516 qRT-F	TGAGTTTTCTGCGTTGTTCG
FRT-Kan F	ATTCCGGGGATCCGTCGAC
FRT-Kan-R	TGTAGGCTGGAGCTGCTTC
ΔVC1142 F1	TGAAAGCGACCATGATGAAA
ΔVC1142 R1	GTCGACGGATCCCCGGAAT TAGTGGATTGATTGCCACACTAGC
ΔVC1142 F2	GAAGCAGCTCCAGCCTACA CATCCCTCATGCATTTTATAACTGATG
ΔVC1142 R2	CACACTGCGCTGAATCACTT
Δvca0722 F1	CACTGGCACTTTCTTCGACA
Δvca0722 R1	GTCGACGGATCCCCGGAAT CATCGCGCTTACCTCATCGTTA
Δvca0722 F2	GAAGCAGCTCCAGCCTACA TAAGCCAAACTGTAAGCTTTTTTATC
Δvca0722 R2	CGGTCTTTGCCCAAATTAAA
Δvca1051 F1	CTTGCGTACCTTGTGGGATT
Δvca1051 R1	GAAGCAGCTCCAGCCTACA TAATACGCTTTAATACCGAACTCAC
Δvca1051 F2	GTCGACGGATCCCCGGAAT CATTACCTTACCTCACGTTGTTG
Δvca1051 R2	TGCTGTGATTGCCTCTATCG
Δvc1863 F1	GAGGTTTTAGGTAGGCAAG
Δvc1863 R1	GTCGACGGATCCCCGGAAT AATAAAAGCGCTTCTTATCACTTCG
Δvc1863 F2	GAAGCAGCTCCAGCCTACA CATACTCCTGTGATTGTTTGTCTTA
Δvc1863 R2	GTCAAATGGGCTACTTCTAACA
Δvc1992 F1	TGTTCGGCTCGGGTAAATAG
Δvc1992 R1	GTCGACGGATCCCCGGAAT CATTGATAACAACTTCCATGCAAAAA
Δvc1992 F2	GAAGCAGCTCCAGCCTACA TAACGAACCCTCAGCAGT
Δvc1992 R2	GCTTGATTGTGACGGGTG
Δvc2492 F1	CAGTGTATCTATCGCATC
Δvc2492 R1	GTCGACGGATCCCCGGAAT CATTGCTTCTTCCTTGAG
Δvc2492 F2	GAAGCAGCTCCAGCCTACA TAATTCAAGGAGAACCCCA
Δvc2492 R2	AATAAAGGGATAGCTCAGG
ΔVC0516 F1	GCAAAGCTTAGTTGCCGAAG
ΔVC0516 R1	GTCGACGGATCCCCGGAAT TAGTAGTAATAAACAATTCTAAGGCC

ΔVC0516 F2	GAAGCAGCTCCAGCCTACA CATGATGTACCACCAAAATAGTC
ΔVC0516 R2	GTTCAAAGTGCGATTGCTGA
SacI VC1863 F	GCGAGCTC ATTTATTAGATGGGCAAATAGAGGC
XbaI VC1863 R	GCTCTAGA TTATTCACCGTAAACGTCATACTTG
VC1863 mut R1	AAAAAGTGGTTATATGGGACAGCATTAATTGCCACAACATAT
VC1863 mut F2	GTTACCCCACTATATGTTGTGGCAATTAATGCTGTCCC
VCA0686 mut sacI F1	GCGAGCTC CGCAAAAATTGATGCAGAGA
VCA0686 mut R1	**T**TC**G**TGCaTg**ca**caaaccttattattga**c**tc**gg** TGAGTCGCACTTAACCTGC
VCA0686 mut F2	ccgagtcaataataaggtttgtgcAtGCACGAA CATGAGCCTTCTCAAGCGCT
VCA0686mut R2 xbaI	Gctctaga CTTCTTCCAACGAAGGATGC
tarB VC1863 mut R1	CTTATAGGTCAGCTTGTAACAATACATGAGG
tarB VC1863 mmut F2	TGTTACAAGCTGACCTATAAGCACTTTTTGG
tarB VCA0686 F2	TTCGTGTATTCATACAAGCTGACGATTAACG ACTTTTTGGCGCA
tarB VCA0686 R1	CGTTAATCGTCAGCTTGTATGAATACACGAA GTAACCGCTCAT
vrrA R t7	TAATACGACTCACTATAGGGAAAACGCCCAAAACAATCTG
vrrA F	ACTGGCCGTCAAATTTGGTT
tarB R	AAAACAAAATCATCTTTCATAACAGC
T7 *tarB* F	TAATACGACTCACTATAGGGCGCCAAAAAGTGCTTAATCG
DtarC F1	TAGCGTCGAGATGACACACC
DtarC R1	GTCGACGGATCCCCGGAAT TTTCACAGAAAGACGCAAAAAAAGA
DtarC F2	GAAGCAGCTCCAGCCTACA AACTGATTTCATAATCTGTATCAAAAAAAAT

DtarC R2	AGGCAAGTGAAGCAACATCA
sacI TarC F	GC GAGCTC GCCTTTACACCAGAGCCAAT
xbaI TarC R	GC TCTAGA CCACCCCAAACATACAAACA

Table 5 Barcoded adapters and sRNA cloning linkers for constructing Illumina sRNA-seq libraries [151]

Name	Sequence	barcode
Linker1	rAppCTGTAGGCACCATCAAT/3ddC/	
Linker2	rAppCACTCGGGCACCAAGGA/3ddC/	
Linker3	rAppTTTAACCGCGAATTCCAG/3ddC/	
BC1a	AATGATACGGCGACCACCGAGATCTACACTCTTTCCCTAC ACGACGCTCTTCCGATCTAACCT	AACC
BC1b	P-GGTTAGATCGGAAGAGCGGTTCAGCAGGAATGCCGAGAC CGATCTCGTATGCCGTCTTCTGCTTG	AACC
BC2a	AATGATACGGCGACCACCGAGATCTACACTCTTTCCCTAC ACGACGCTCTTCCGATCTTTGGT	TTGG
BC2b	P-CCAAAGATCGGAAGAGCGGTTCAGCAGGAATGCCGAGAC CGATCTCGTATGCCGTCTTCTGCTTG	TTGG
BC3a	AATGATACGGCGACCACCGAGATCTACACTCTTTCCCTAC ACGACGCTCTTCCGATCTCCAAT	CCAA
BC3b	P-TTGGAGATCGGAAGAGCGGTTCAGCAGGAATGCCGAGAC CGATCTCGTATGCCGTCTTCTGCTTG	CCAA

Table 6 Plasmids used in this study

Name	Description	Reference
pBAD24	pBR322 based plasmid with arabinose inducible promoter and ampR selectable marker	[184]
pToxT	pBAD24 backbone with toxT ORF and RBS cloned into the NcoI/XbaI site	This study
pToxT (ΔHLH)	pBAD24 back bone with toxT ΔHLH ORF and RBS loned into the NcoI/XbaI site	This study
pMMB 67EH	Apr, IncQ broad-host-range cloning vector	[191]

p*tarB*	pMMB67EH with fragment containing the tarB sequence, promoter and terminator cloned into the SacI/XbaI sites	This study
p*tarB**	pMMB67EH with fragment containing the tarB sequence, promoter and terminator cloned into the SacI/XbaI sites	This study
pMAL-TEV-His-thr-*toxT*	pMAL vector from NEB modified to contain a TEV protease cleavage site, a 6x his tag and a thrombin cleavage site in between MBP and toxT, the toxT ORF was cloned into the NdeI/BamHI site of this vector	This study
p*tarA*	Topo vector pCR 2.1 containing the tarA sequence including promoter and previously described toxboxes	This study
pGEM-T	Empty vector containing the colE1 origin, used as no insert control for competitions against ptarA containing strain	Promega
pGP704 cm tcpF C-FLAG	Derivitive of the pGP704 plasmid containing the CAT gene swapped for bla and the C-terminal 300 base pairs of the tcpF ORF with in-frame FLAG tag cloned into the EcoRV/KpnI sites. Used for generating the C-Terminal FLAG fusion to TcpF	This study
ptarB-300	Topo vector pCR2.1 containing the +3 to -380bp of the tarB promoter	This study
pTarB-GFP(ASV)	Derivitive of the pGP704 plasmid containing the CAT gene swapped for bla and the -380 to +3bp of the	This study

	TarB promoter cloned ahead of the Gfp(ASV) [60, 158] allele in the SmaI site	
pJML01	Modified pBAD24 back bone altered to make transcription start at the NheI site	[109]
pJML01-*tarB*	Modified pBAD24 back bone with the *tarB* sRNA or various mutations made in in the *tarB* sequence cloned into the NheI/SphI site	This study
pJML01-*tarB1863mut*	Mutant allele of *tarB* cloned into the into the NheI/SphI site of pJML01 to restore interaction with the mutated VC1863 allele	This study
pJML01-*tarB0686mut*	Mutant allele of *tarB* cloned into the into the NheI/SphI site of pJML01 to restore interaction with the mutated VCA0686 allele	Promega
p1863 comp	pMMB67EH with fragment the VC1863 open reading frame and upstream sequence into the SacI/XbaI sites	This study
pBR-flp	A pBR322 based plasmid containing the flp recombinates gene under the control of the λ p_R promoter and a tempreture sensitive allele of the lambda repressor (λ cI857)	This study
pCVD442-lac	A deriviative of the pCVD442 allelic exchange vector carrying the *sacB* gene for sucrose counter selection and the *lacZ* gene and MCS from the pUC19 vector	[185]

pTarC	A derivative of the pMMB plasmid with the sequence of the putative TarC sRNA sequence, promoter included	This study

Table 7 strains

Name	Genotype	Source
AC53	Wiltype E7946 El Tor, O1 Ogawa	Laboratory Collection
AC3745	Δ*lacZ*	This study
AC3763	Δ*toxT*	This study
AC3744	Δ*tarB*	This study
AC3748	Δ*tarB*Δ*lacZ*	This study
AC3746	Δ*tarA*	This study
AC3749	Δ*tarA*Δ*lacZ*	This study
AC3757	Δ*tarB*Δ*toxT*	This study
AC3794	Δ*tarBtcpF**	This study
AC3795	Δ*tarB*Δ*lacZtcpF**	This study
AC3765	Δ*hfq*	This study
AC468	Δ*toxR* (ΔHLH) C6709 El Tor O1 Inaba	Laboratory Collection
AC522	Δ*tcpPH*::*KanR* res-tet-res C6709 El Tor O1 Inaba	Laboratory Collection
AC3780	*tarB**	This study
AC3781	*tcpF**	This study
AC3782	*tarB***tcpF**	This study
AC4099	ΔVC1142::FRT	This study
AC4125	ΔVCA0722::FRT	This study
AC4127	ΔVC1992::FRT	This study
AC4128	ΔVC2492::KanR	This study
AC4145	ΔVC0516::FRT	This study
AC4138	ΔVCA0722::FRT Δ*lacZ*	This study
AC4140	ΔVC1863::FRT Δ*lacZ*	This study

AC4148	VCA1051::KanR	This study
AC4158	Δ*tarB* VC1863mut	This study
AC4171	Δ*tarB* VCA0686mut	This study
AC4092	Δ*tarC*::FRT	This study
AC4131	Δ*tarC*::FRT Δ*lacZ*	This study

All strains, except where noted, are derivatives of wildtype E7946

References

1. Koch R (1883) An address on cholera and its bacillus. BMJ 2: 403–407.
2. Seas C, Gotuzzo E (2009) Mandell: Mandell, Douglas, and Bennett's Principles and Practice of Infectious Diseases. 8 p.
3. Waldor MK, Mekalanos JJ (1994) Emergence of a new cholera pandemic: molecular analysis of virulence determinants in Vibrio cholerae O139 and development of a live vaccine prototype. The Journal of infectious diseases 170: 278-283.
4. Nelson EJ, Harris JB, Morris JG, Calderwood SB, Camilli A (2009) Cholera transmission: the host, pathogen and bacteriophage dynamic. Nat Rev Micro 7: 693-702.
5. Waldor MK, Mekalanos JJ (1996) Lysogenic conversion by a filamentous phage encoding cholera toxin. Science 272: 1910-1914.
6. Bishop AL, Camilli A (2011) Vibrio cholerae: lessons for mucosal vaccine design. Expert review of vaccines 10: 79-94.
7. Safa A, Nair GB, Kong RY (2010) Evolution of new variants of Vibrio cholerae O1. Trends in microbiology 18: 46-54.
8. Wachsmuth K, Blake P, Olsvik O (1994) Vibrio cholerae and Cholera. Washington, D.C.: ASM Press.
9. Almagro-Moreno S, Boyd EF (2009) Sialic acid catabolism confers a competitive advantage to pathogenic vibrio cholerae in the mouse intestine. Infection and immunity 77: 3807-3816.
10. Dziejman M, Balon E, Boyd D, Fraser CM, Heidelberg JF, et al. (2002) Comparative genomic analysis of Vibrio cholerae: genes that correlate with cholera endemic and pandemic disease. Proceedings of the National Academy of Sciences of the United States of America 99: 1556-1561.
11. Murphy RA, Boyd EF (2008) Three pathogenicity islands of Vibrio cholerae can excise from the chromosome and form circular intermediates. Journal of Bacteriology 190: 636-647.
12. Davies BW, Bogard RW, Young TS, Mekalanos JJ (2012) Coordinated Regulation of Accessory Genetic Elements Produces Cyclic Di-Nucleotides for V. cholerae Virulence. Cell 149: 358-370.

13. Greenough DBaWB (1992) Cholera. New York: Plenum Publishing Corporation. 372 p.
14. Murley YM, Behari J, Griffin R, Calderwood SB (2000) Classical and El Tor biotypes of Vibrio cholerae differ in timing of transcription of tcpPH during growth in inducing conditions. Infect Immun 68: 3010-3014.
15. Mukhopadhyay AK, Garg S, Saha PK, Takeda Y, Bhattacharya SK, et al. (1996) Comparative analysis of factors promoting optimal production of cholera toxin by Vibrio cholerae O1 (classical & E1Tor biotypes) & O139. Indian J Med Res 104: 129-133.
16. DiRita VJ, Neely M, Taylor RK, Bruss PM (1996) Differential expression of the ToxR regulon in classical and E1 Tor biotypes of Vibrio cholerae is due to biotype-specific control over toxT expression. Proc Natl Acad Sci U S A 93: 7991-7995.
17. Skorupski K, Taylor RK (1997) Control of the ToxR virulence regulon in Vibrio cholerae by environmental stimuli. Molecular microbiology 25: 1003-1009.
18. Bhattacharya SK (1995) Cholera outbreaks: role of oral rehydration therapy. Journal of the Indian Medical Association 93: 237-238.
19. Guerrant RL, Carneiro-Filho BA, Dillingham RA (2003) Cholera, diarrhea, and oral rehydration therapy: triumph and indictment. Clinical infectious diseases : an official publication of the Infectious Diseases Society of America 37: 398-405.
20. Wozniak RA, Fouts DE, Spagnoletti M, Colombo MM, Ceccarelli D, et al. (2009) Comparative ICE genomics: insights into the evolution of the SXT/R391 family of ICEs. PLoS genetics 5: e1000786.
21. Schild S, Nelson EJ, Bishop AL, Camilli A (2009) Characterization of Vibrio cholerae outer membrane vesicles as a candidate vaccine for cholera. Infection and immunity 77: 472-484.
22. Date KA, Vicari A, Hyde TB, Mintz E, Danovaro-Holliday MC, et al. (2011) Considerations for oral cholera vaccine use during outbreak after earthquake in Haiti, 2010-2011. Emerging infectious diseases 17: 2105-2112.
23. Calain P, Chaine JP, Johnson E, Hawley ML, O'Leary MJ, et al. (2004) Can oral cholera vaccination play a role in controlling a cholera outbreak? Vaccine 22: 2444-2451.
24. Miller CJ, Feachem RG, Drasar BS (1985) Cholera epidemiology in developed and developing countries: new thoughts on transmission, seasonality, and control. Lancet 1: 261-262.
25. Colwell RR, Huq A (1994) Environmental reservoir of Vibrio cholerae. The causative agent of cholera. Annals of the New York Academy of Sciences 740: 44-54.
26. Meibom KL, Li XB, Nielsen AT, Wu CY, Roseman S, et al. (2004) The Vibrio cholerae chitin utilization program. Proceedings of the National Academy of Sciences of the United States of America 101: 2524-2529.
27. Alam M, Sultana M, Nair GB, Siddique AK, Hasan NA, et al. (2007) Viable but nonculturable Vibrio cholerae O1 in biofilms in the aquatic environment and their role in cholera transmission. Proceedings of the National Academy of Sciences of the United States of America 104: 17801-17806.
28. Miller CJ, Drasar BS, Feachem RG (1982) Cholera and estuarine salinity in Calcutta and London. Lancet 1: 1216-1218.

29. Lara RJ, Neogi SB, Islam MS, Mahmud ZH, Yamasaki S, et al. (2009) Influence of catastrophic climatic events and human waste on Vibrio distribution in the Karnaphuli estuary, Bangladesh. EcoHealth 6: 279-286.
30. Meibom KL, Blokesch M, Dolganov NA, Wu CY, Schoolnik GK (2005) Chitin induces natural competence in Vibrio cholerae. Science 310: 1824-1827.
31. Blokesch M, Schoolnik GK (2007) Serogroup conversion of Vibrio cholerae in aquatic reservoirs. PLoS pathogens 3: e81.
32. Taylor RK, Miller VL, Furlong DB, Mekalanos JJ (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.
33. Klose KE (2000) The suckling mouse model of cholera. Trends in microbiology 8: 189-191.
34. Finkelstein RA, LoSpalluto JJ (1969) Pathogenesis of experimental cholera. Preparation and isolation of choleragen and choleragenoid. The Journal of experimental medicine 130: 185-202.
35. Finkelstein RA, LoSpalluto JJ (1970) Production of highly purified choleragen and choleragenoid. The Journal of infectious diseases 121: Suppl 121:163+.
36. Kirn TJ, Jude BA, Taylor RK (2005) A colonization factor links Vibrio cholerae environmental survival and human infection. Nature 438: 863-866.
37. Parsot C, Mekalanos JJ (1992) Structural analysis of the acfA and acfD genes of Vibrio cholerae: effects of DNA topology and transcriptional activators on expression. J Bacteriol 174: 5211-5218.
38. Merrell DS, Hava DL, Camilli A (2002) Identification of novel factors involved in colonization and acid tolerance of Vibrio cholerae. Mol Microbiol 43: 1471-1491.
39. Craig L, Pique ME, Tainer JA (2004) Type IV pilus structure and bacterial pathogenicity. Nature reviews Microbiology 2: 363-378.
40. Giron JA, Gomez-Duarte OG, Jarvis KG, Kaper JB (1997) Longus pilus of enterotoxigenic Escherichia coli and its relatedness to other type-4 pili--a minireview. Gene 192: 39-43.
41. Donnenberg MS, Zhang HZ, Stone KD (1997) Biogenesis of the bundle-forming pilus of enteropathogenic Escherichia coli: reconstitution of fimbriae in recombinant E. coli and role of DsbA in pilin stability--a review. Gene 192: 33-38.
42. Kirn TJ, Taylor RK (2005) TcpF is a soluble colonization factor and protective antigen secreted by El Tor and classical O1 and O139 Vibrio cholerae serogroups. Infect Immun 73: 4461-4470.
43. Faruque SM, Zhu J, Asadulghani, Kamruzzaman M, Mekalanos JJ (2003) 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.
44. Faruque SM, Asadulghani, Saha MN, Alim AR, Albert MJ, et al. (1998) Analysis of clinical and environmental strains of nontoxigenic Vibrio cholerae for susceptibility to CTXPhi: molecular basis for origination of new strains with epidemic potential. Infect Immun 66: 5819-5825.

45. Pratt JT, Ismail AM, Camilli A (2010) PhoB regulates both environmental and virulence gene expression in Vibrio cholerae. Molecular microbiology 77: 1595-1605.
46. Kirn TJ, Bose N, Taylor RK (2003) Secretion of a soluble colonization factor by the TCP type 4 pilus biogenesis pathway in Vibrio cholerae. Mol Microbiol 49: 81-92.
47. Krebs SJ, Kirn TJ, Taylor RK (2009) Genetic mapping of secretion and functional determinants of the Vibrio cholerae TcpF colonization factor. Journal of bacteriology 191: 3665-3676.
48. Vanden Broeck D, Horvath C, De Wolf MJ (2007) Vibrio cholerae: cholera toxin. The international journal of biochemistry & cell biology 39: 1771-1775.
49. Nielsen AT, Dolganov NA, Otto G, Miller MC, Wu CY, et al. (2006) RpoS controls the Vibrio cholerae mucosal escape response. PLoS pathogens 2: e109.
50. McCardell BA, Madden JM, Shah DB (1985) Isolation and characterization of a cytolysin produced by Vibrio cholerae serogroup non-O1. Canadian journal of microbiology 31: 711-720.
51. Trucksis M, Galen JE, Michalski J, Fasano A, Kaper JB (1993) Accessory cholera enterotoxin (Ace), the third toxin of a Vibrio cholerae virulence cassette. Proceedings of the National Academy of Sciences of the United States of America 90: 5267-5271.
52. Fasano A, Baudry B, Pumplin DW, Wasserman SS, Tall BD, et al. (1991) Vibrio cholerae produces a second enterotoxin, which affects intestinal tight junctions. Proceedings of the National Academy of Sciences of the United States of America 88: 5242-5246.
53. Levine MM, Kaper JB, Herrington D, Losonsky G, Morris JG, et al. (1988) Volunteer studies of deletion mutants of Vibrio cholerae O1 prepared by recombinant techniques. Infection and immunity 56: 161-167.
54. Jude BA, Martinez RM, Skorupski K, Taylor RK (2009) Levels of the secreted Vibrio cholerae attachment factor GbpA are modulated by quorum-sensing-induced proteolysis. Journal of bacteriology 191: 6911-6917.
55. Peterson KM (2002) Expression of Vibrio cholerae virulence genes in response to environmental signals. Current issues in intestinal microbiology 3: 29-38.
56. Childers BM, Klose KE (2007) Regulation of virulence in Vibrio cholerae: the ToxR regulon. Future Microbiol 2: 335-344.
57. Lee SH, Hava DL, Waldor MK, Camilli A (1999) Regulation and temporal expression patterns of Vibrio cholerae virulence genes during infection. Cell 99: 625-634.
58. Schild S, Tamayo R, Nelson EJ, Qadri F, Calderwood SB, et al. (2007) Genes induced late in infection increase fitness of Vibrio cholerae after release into the environment. Cell Host & Microbe 2: 264-277.
59. Abuaita BH, Withey JH (2011) Termination of Vibrio cholerae virulence gene expression is mediated by proteolysis of the major virulence activator, ToxT. Molecular microbiology 81: 1640-1653.
60. Nielsen AT, Dolganov NA, Rasmussen T, Otto G, Miller MC, et al. A bistable switch and anatomical site control Vibrio cholerae virulence gene expression in the intestine. PLoS Pathog 6.

61. Lin W, Kovacikova G, Skorupski K (2007) The quorum sensing regulator HapR downregulates the expression of the virulence gene transcription factor AphA in Vibrio cholerae by antagonizing Lrp- and VpsR-mediated activation. Molecular microbiology 64: 953-967.
62. Merrell DS, Butler SM, Qadri F, Dolganov NA, Alam A, et al. (2002) Host-induced epidemic spread of the cholera bacterium. Nature 417: 642-645.
63. Hartley DM, Morris JG, Jr., Smith DL (2006) Hyperinfectivity: a critical element in the ability of V. cholerae to cause epidemics? PLoS medicine 3: e7.
64. Alam A, Larocque RC, Harris JB, Vanderspurt C, Ryan ET, et al. (2005) Hyperinfectivity of human-passaged Vibrio cholerae can be modeled by growth in the infant mouse. Infection and immunity 73: 6674-6679.
65. Butler SM, 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.
66. Butler SM, Nelson EJ, Chowdhury N, Faruque SM, Calderwood SB, et al. (2006) Cholera stool bacteria repress chemotaxis to increase infectivity. Molecular microbiology 60: 417-426.
67. Karaolis DK, Somara S, Maneval DR, Johnson JA, Kaper JB (1999) A bacteriophage encoding a pathogenicity island, a type-IV pilus and a phage receptor in cholera bacteria. Nature 399: 375-379.
68. Higgins DE, Nazareno E, DiRita VJ (1992) The virulence gene activator ToxT from Vibrio cholerae is a member of the AraC family of transcriptional activators. J Bacteriol 174: 6974-6980.
69. Yu RR, DiRita VJ (1999) Analysis of an autoregulatory loop controlling ToxT, cholera toxin, and toxin-coregulated pilus production in Vibrio cholerae. J Bacteriol 181: 2584-2592.
70. Medrano AI, DiRita VJ, Castillo G, Sanchez J (1999) Transient transcriptional activation of the Vibrio cholerae El Tor virulence regulator toxT in response to culture conditions. Infect Immun 67: 2178-2183.
71. Kovacikova G, Skorupski K (2001) Overlapping binding sites for the virulence gene regulators AphA, AphB and cAMP-CRP at the Vibrio cholerae tcpPH promoter. Mol Microbiol 41: 393-407.
72. Kovacikova G, Lin W, Skorupski K (2010) The LysR-type virulence activator AphB regulates the expression of genes in Vibrio cholerae in response to low pH and anaerobiosis. Journal of Bacteriology 192: 4181-4191.
73. Goss TJ, Seaborn CP, Gray MD, Krukonis ES (2010) Identification of the TcpP-binding site in the toxT promoter of Vibrio cholerae and the role of ToxR in TcpP-mediated activation. Infection and immunity 78: 4122-4133.
74. Krukonis ES, Yu RR, Dirita VJ (2000) The Vibrio cholerae ToxR/TcpP/ToxT virulence cascade: distinct roles for two membrane-localized transcriptional activators on a single promoter. Molecular microbiology 38: 67-84.
75. Marrero K, Sanchez A, Rodriguez-Ulloa A, Gonzalez LJ, Castellanos-Serra L, et al. (2009) Anaerobic growth promotes synthesis of colonization factors encoded at the Vibrio pathogenicity island in Vibrio cholerae El Tor. Res Microbiol 160: 48-56.

76. Iwanaga M, Yamamoto K, Higa N, Ichinose Y, Nakasone N, et al. (1986) Culture conditions for stimulating cholera toxin production by Vibrio cholerae O1 El Tor. Microbiol Immunol 30: 1075-1083.
77. Abuaita BH, Withey JH (2009) Bicarbonate Induces Vibrio cholerae virulence gene expression by enhancing ToxT activity. Infect Immun 77: 4111-4120.
78. Zhu J, Miller MB, Vance RE, Dziejman M, Bassler BL, et al. (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.
79. Kovacikova G, Skorupski K (2002) Regulation of virulence gene expression in Vibrio cholerae by quorum sensing: HapR functions at the aphA promoter. Molecular microbiology 46: 1135-1147.
80. Waters CM, Lu W, Rabinowitz JD, Bassler BL (2008) Quorum sensing controls biofilm formation in Vibrio cholerae through modulation of cyclic di-GMP levels and repression of vpsT. Journal of bacteriology 190: 2527-2536.
81. Kovacikova G, Skorupski K (2001) Overlapping binding sites for the virulence gene regulators AphA, AphB and cAMP-CRP at the Vibrio cholerae tcpPH promoter. Molecular microbiology 41: 393-407.
82. Liang W, Pascual-Montano A, Silva AJ, Benitez JA (2007) The cyclic AMP receptor protein modulates quorum sensing, motility and multiple genes that affect intestinal colonization in Vibrio cholerae. Microbiology 153: 2964-2975.
83. Tischler AD, Camilli A (2005) Cyclic diguanylate regulates Vibrio cholerae virulence gene expression. Infection and immunity 73: 5873-5882.
84. Cotter PA, Stibitz S (2007) c-di-GMP-mediated regulation of virulence and biofilm formation. Current opinion in microbiology 10: 17-23.
85. Beyhan S, Tischler AD, Camilli A, Yildiz FH (2006) Transcriptome and phenotypic responses of Vibrio cholerae to increased cyclic di-GMP level. Journal of bacteriology 188: 3600-3613.
86. Tamayo R, Tischler AD, Camilli A (2005) The EAL domain protein VieA is a cyclic diguanylate phosphodiesterase. The Journal of biological chemistry 280: 33324-33330.
87. Provenzano D, Lauriano CM, Klose KE (2001) Characterization of the role of the ToxR-modulated outer membrane porins OmpU and OmpT in Vibrio cholerae virulence. Journal of bacteriology 183: 3652-3662.
88. Provenzano D, Klose KE (2000) Altered expression of the ToxR-regulated porins OmpU and OmpT diminishes Vibrio cholerae bile resistance, virulence factor expression, and intestinal colonization. Proceedings of the National Academy of Sciences of the United States of America 97: 10220-10224.
89. Withey JH, Dirita VJ (2006) The toxbox: specific DNA sequence requirements for activation of Vibrio cholerae virulence genes by ToxT. Mol Microbiol 59: 1779-1789.
90. Childers BM, Weber GG, Prouty MG, Castaneda MM, Peng F, et al. (2007) Identification of residues critical for the function of the Vibrio cholerae virulence regulator ToxT by scanning alanine mutagenesis. J Mol Biol 367: 1413-1430.
91. Bellair M, Withey JH (2008) Flexibility of Vibrio cholerae ToxT in transcription activation of genes having altered promoter spacing. J Bacteriol 190: 7925-7931.

92. Bradley ES, Bodi K, Ismail AM, Camilli A (2011) A genome-wide approach to discovery of small RNAs involved in regulation of virulence in Vibrio cholerae. PLoS pathogens 7: e1002126.
93. Weber GG, Klose KE (2011) The complexity of ToxT-dependent transcription in Vibrio cholerae. The Indian journal of medical research 133: 201-206.
94. Krishnan HH, Ghosh A, Paul K, Chowdhury R (2004) Effect of anaerobiosis on expression of virulence factors in Vibrio cholerae. Infect Immun 72: 3961-3967.
95. Navarre WW, Porwollik S, Wang Y, McClelland M, Rosen H, et al. (2006) Selective silencing of foreign DNA with low GC content by the H-NS protein in Salmonella. Science 313: 236-238.
96. Lowden MJ, Skorupski K, Pellegrini M, Chiorazzo MG, Taylor RK, et al. (2010) Structure of Vibrio cholerae ToxT reveals a mechanism for fatty acid regulation of virulence genes. Proc Natl Acad Sci USA 107: 2860-2865.
97. Shakhnovich EA, Hung DT, Pierson E, Lee K, Mekalanos JJ (2007) Virstatin inhibits dimerization of the transcriptional activator ToxT. Proceedings of the National Academy of Sciences of the United States of America 104: 2372-2377.
98. Hsiao A, Xu X, Kan B, Kulkarni RV, Zhu J (2009) Direct regulation by the Vibrio cholerae regulator ToxT to modulate colonization and anticolonization pilus expression. INFECTION AND IMMUNITY 77: 1383-1388.
99. Peterson KM, Mekalanos JJ (1988) Characterization of the Vibrio cholerae ToxR regulon: identification of novel genes involved in intestinal colonization. Infect Immun 56: 2822-2829.
100. Altuvia S (2007) Identification of bacterial small non-coding RNAs: experimental approaches. Current Opinion in Microbiology 10: 257-261.
101. Fozo EM, Makarova KS, Shabalina SA, Yutin N, Koonin EV, et al. (2010) Abundance of type I toxin-antitoxin systems in bacteria: searches for new candidates and discovery of novel families. Nucleic acids research 38: 3743-3759.
102. Vogel J (2009) A rough guide to the non-coding RNA world of Salmonella. Mol Microbiol 71: 1-11.
103. Hammer BK, Bassler BL (2007) Regulatory small RNAs circumvent the conventional
quorum sensing pathway in pandemic Vibrio cholerae. PNAS: 5.
104. Song T, Wai SN (2009) A novel sRNA that modulates virulence and environmental fitness of Vibrio cholerae. RNA biology 6: 254-258.
105. Loh E, Dussurget O, Gripenland J, Vaitkevicius K, Tiensuu T, et al. (2009) A trans-acting riboswitch controls expression of the virulence regulator PrfA in Listeria monocytogenes. Cell 139: 770-779.
106. Richards GR, Vanderpool CK (2011) Molecular call and response: the physiology of bacterial small RNAs. Biochimica et biophysica acta 1809: 525-531.
107. Opdyke JA, Kang JG, Storz G (2004) GadY, a small-RNA regulator of acid response genes in Escherichia coli. Journal of bacteriology 186: 6698-6705.
108. Liu JM, Camilli A (2009) A broadening world of bacterial small RNAs. Curr Opin Microbiol.
109. Viegas SC, Arraiano CM (2008) Regulating the regulators: How ribonucleases dictate the rules in the control of small non-coding RNAs. RNA biology 5: 230-243.

110. Rutherford ST, van Kessel JC, Shao Y, Bassler BL (2011) AphA and LuxR/HapR reciprocally control quorum sensing in vibrios. Genes & development 25: 397-408.
111. Song T, Sabharwal D, Wai SN (2010) VrrA mediates Hfq-dependent regulation of OmpT synthesis in Vibrio cholerae. Journal of molecular biology 400: 682-688.
112. Aiba H (2007) Mechanism of RNA silencing by Hfq-binding small RNAs. Current Opinion in Microbiology 10: 134-139.
113. Kajitani M, Ishihama A (1991) Identification and sequence determination of the host factor gene for bacteriophage Q beta. Nucleic acids research 19: 1063-1066.
114. Liu JM, Livny J, Lawrence MS, Kimball MD, Waldor MK, et al. (2009) Experimental discovery of sRNAs in Vibrio cholerae by direct cloning, 5S/tRNA depletion and parallel sequencing. Nucleic Acids Res 37: e46.
115. Storz G, Vogel J, Wassarman KM (2011) Regulation by small RNAs in bacteria: expanding frontiers. Molecular cell 43: 880-891.
116. Papenfort K, Bouvier M, Mika F, Sharma CM, Vogel J (2010) Evidence for an autonomous 5' target recognition domain in an Hfq-associated small RNA. Proceedings of the National Academy of Sciences of the United States of America 107: 20435-20440.
117. Frohlich KS, Vogel J (2009) Activation of gene expression by small RNA. Curr Opin Microbiol 12: 674-682.
118. Pfeiffer V, Papenfort K, Lucchini S, Hinton JC, Vogel J (2009) Coding sequence targeting by MicC RNA reveals bacterial mRNA silencing downstream of translational initiation. Nat Struct Mol Biol 16: 840-846.
119. Papenfort K, Said N, Welsink T, Lucchini S, Hinton JC, et al. (2009) Specific and pleiotropic patterns of mRNA regulation by ArcZ, a conserved, Hfq-dependent small RNA. Mol Microbiol 74: 139-158.
120. Babitzke P, Romeo T (2007) CsrB sRNA family: sequestration of RNA-binding regulatory proteins. Current Opinion in Microbiology 10: 156-163.
121. Lenz DH, Miller MB, Zhu J, Kulkarni RV, Bassler BL (2005) CsrA and three redundant small RNAs regulate quorum sensing in Vibrio cholerae. Molecular microbiology 58: 1186-1202.
122. Yakhnin H, Pandit P, Petty TJ, Baker CS, Romeo T, et al. (2007) CsrA of Bacillus subtilis regulates translation initiation of the gene encoding the flagellin protein (hag) by blocking ribosome binding. Mol Microbiol 64: 1605-1620.
123. Jonas K, Edwards AN, Ahmad I, Romeo T, Romling U, et al. (2009) Complex regulatory network encompassing the Csr, c-di-GMP and motility systems of Salmonella Typhimurium. Environ Microbiol.
124. Liu MY, Gui G, Wei B, Preston JF, 3rd, Oakford L, et al. (1997) The RNA molecule CsrB binds to the global regulatory protein CsrA and antagonizes its activity in Escherichia coli. J Biol Chem 272: 17502-17510.
125. Wassarman K (2007) 6S RNA: a small RNA regulator of transcription. Current Opinion in Microbiology 10: 164-168.
126. Vogel J, Wagner E (2007) Target identification of small noncoding RNAs in bacteria. Current Opinion in Microbiology 10: 262-270.
127. Valentin-Hansen P, Eriksen M, Udesen C (2004) The bacterial Sm-like protein Hfq: a key player in RNA transactions. Mol Microbiol 51: 1525-1533.

128. Nielsen JS, Boggild A, Andersen CB, Nielsen G, Boysen A, et al. (2007) An Hfq-like protein in archaea: crystal structure and functional characterization of the Sm protein from Methanococcus jannaschii. RNA 13: 2213-2223.
129. Link TM, Valentin-Hansen P, Brennan RG (2009) Structure of Escherichia coli Hfq bound to polyriboadenylate RNA. Proc Natl Acad Sci U S A 106: 19292-19297.
130. Maki K, Uno K, Morita T, Aiba H (2008) RNA, but not protein partners, is directly responsible for translational silencing by a bacterial Hfq-binding small RNA. Proceedings of the National Academy of Sciences of the United States of America 105: 10332-10337.
131. Sittka A, Lucchini S, Papenfort K, Sharma CM, Rolle K, et al. (2008) Deep sequencing analysis of small noncoding RNA and mRNA targets of the global post-transcriptional regulator, Hfq. PLoS genetics 4: e1000163.
132. Christiansen JK, Nielsen JS, Ebersbach T, Valentin-Hansen P, Sogaard-Andersen L, et al. (2006) Identification of small Hfq-binding RNAs in Listeria monocytogenes. RNA 12: 1383-1396.
133. Ding Y, Davis BM, Waldor MK (2004) Hfq is essential for Vibrio cholerae virulence and downregulates sigma expression. Molecular microbiology 53: 345-354.
134. Moon K, Gottesman S (2011) Competition among Hfq-binding small RNAs in Escherichia coli. Molecular microbiology 82: 1545-1562.
135. Davis BM, Waldor MK (2007) RNase E-dependent processing stabilizes MicX, a Vibrio cholerae sRNA. Mol Microbiol 65: 373-385.
136. Urban JH, Vogel J (2007) Translational control and target recognition by Escherichia coli small RNAs in vivo. Nucleic Acids Res 35: 1018-1037.
137. Pfeiffer V, Sittka A, Tomer R, Tedin K, Brinkmann V, et al. (2007) A small non-coding RNA of the invasion gene island (SPI-1) represses outer membrane protein synthesis from the Salmonella core genome. Mol Microbiol 66: 1174-1191.
138. Geisinger E, Adhikari RP, Jin R, Ross HF, Novick RP (2006) Inhibition of rot translation by RNAIII, a key feature of agr function. Molecular microbiology 61: 1038-1048.
139. Novick RP, Geisinger E (2008) Quorum sensing in staphylococci. Annual review of genetics 42: 541-564.
140. Gong H, Vu GP, Bai Y, Chan E, Wu R, et al. (2011) A Salmonella small non-coding RNA facilitates bacterial invasion and intracellular replication by modulating the expression of virulence factors. PLoS pathogens 7: e1002120.
141. Padalon-Brauch G, Hershberg R, Elgrably-Weiss M, Baruch K, Rosenshine I, et al. (2008) Small RNAs encoded within genetic islands of Salmonella typhimurium show host-induced expression and role in virulence. Nucleic acids research 36: 1913-1927.
142. Pfeiffer V, Sittka A, Tomer R, Tedin K, Brinkmann V, et al. (2007) A small non-coding RNA of the invasion gene island (SPI-1) represses outer membrane protein synthesis from the Salmonella core genome. Molecular microbiology 66: 1174-1191.

143. Song T, Mika F, Lindmark B, Liu Z, Schild S, et al. (2008) A new Vibrio choleraesRNA modulates colonization and affects release of outer membrane vesicles. Mol Microbiol 70: 100-111.
144. Richard AL, Withey JH, Beyhan S, Yildiz F, Dirita VJ (2010) The Vibrio cholerae virulence regulatory cascade controls glucose uptake through activation of TarA, a small regulatory RNA. Mol Microbiol 78: 1171-1181.
145. Sittka A, Vogel J (2008) A Glimpse at the Evolution of Virulence Control. Cell Host & Microbe 4: 310-312.
146. Boisset S, Geissmann T, Huntzinger E, Fechter P, Bendridi N, et al. (2007) Staphylococcus aureus RNAIII coordinately represses the synthesis of virulence factors and the transcription regulator Rot by an antisense mechanism. Genes & development 21: 1353-1366.
147. Liu Y, Wu N, Dong J, Gao Y, Zhang X, et al. (2010) Hfq is a global regulator that controls the pathogenicity of Staphylococcus aureus. PloS one 5.
148. Coornaert A, Lu A, Mandin P, Springer M, Gottesman S, et al. (2010) MicA sRNA links the PhoP regulon to cell envelope stress. Molecular microbiology 76: 467-479.
149. Yu RR, Dirita VJ (2002) Regulation of gene expression in Vibrio choleraeby ToxT involves both antirepression and RNA polymerase stimulation. Mol Microbiol 43: 119-134.
150. Stonehouse EA, Hulbert RR, Nye MB, Skorupski KA, Taylor RK (2010) H-NS binding and repression of the ctx promoter in Vibrio cholerae. Journal of Bacteriology.
151. Majerczyk CD, Dunman PM, Luong TT, Lee CY, Sadykov MR, et al. Direct targets of CodY in Staphylococcus aureus. J Bacteriol 192: 2861-2877.
152. Withey JH, DiRita VJ (2006) The toxbox: specific DNA sequence requirements for activation of Vibrio cholerae virulence genes by ToxT. Mol Microbiol 59: 1779-1789.
153. Tjaden B (2008) TargetRNA: a tool for predicting targets of small RNA action in bacteria. Nucleic Acids Res 36: W109-113.
154. Sperandio V, Giron JA, Silveira WD, Kaper JB (1995) The OmpU outer membrane protein, a potential adherence factor of Vibrio cholerae. Infect Immun 63: 4433-4438.
155. Davies BW, Bogard RW, Mekalanos JJ (2011) Mapping the regulon of Vibrio cholerae ferric uptake regulator expands its known network of gene regulation. Proceedings of the National Academy of Sciences of the United States of America.
156. Livny J, Teonadi H, Livny M, Waldor MK (2008) High-throughput, kingdom-wide prediction and annotation of bacterial non-coding RNAs. PLoS One 3: e3197.
157. Selinger DW, Cheung KJ, Mei R, Johansson EM, Richmond CS, et al. (2000) RNA expression analysis using a 30 base pair resolution Escherichia coli genome array. Nature biotechnology 18: 1262-1268.
158. Andersen JB, Sternberg C, Poulsen LK, Bjorn SP, Givskov M, et al. (1998) New unstable variants of green fluorescent protein for studies of transient gene expression in bacteria. Appl Environ Microbiol 64: 2240-2246.

159. Higgins DE, DiRita VJ (1994) Transcriptional control of toxT, a regulatory gene in the ToxR regulon of Vibrio cholerae. Mol Microbiol 14: 17-29.
160. Krukonis ES, Yu RR, Dirita VJ (2000) The Vibrio cholerae ToxR/TcpP/ToxT virulence cascade: distinct roles for two membrane-localized transcriptional activators on a single promoter. Mol Microbiol 38: 67-84.
161. Iwanaga M, Yamamoto K (1985) New medium for the production of cholera toxin by Vibrio cholerae O1 biotype El Tor. J Clin Microbiol 22: 405-408.
162. Liu Z, Yang M, Peterfreund GL, Tsou AM, Selamoglu N, et al. (2011) Vibrio cholerae anaerobic induction of virulence gene expression is controlled by thiol-based switches of virulence regulator AphB. Proceedings of the National Academy of Sciences of the United States of America 108: 810-815.
163. Kovacikova G, Skorupski K (1999) A Vibrio cholerae LysR homolog, AphB, cooperates with AphA at the tcpPH promoter to activate expression of the ToxR virulence cascade. Journal of Bacteriology 181: 4250-4256.
164. Crawford JA, Kaper JB, DiRita VJ (1998) Analysis of ToxR-dependent transcription activation of ompU, the gene encoding a major envelope protein in Vibrio cholerae. Mol Microbiol 29: 235-246.
165. Sezonov G, Joseleau-Petit D, D'Ari R (2007) Escherichia coli physiology in Luria-Bertani broth. Journal of bacteriology 189: 8746-8749.
166. Sengupta N, Paul K, Chowdhury R (2003) The global regulator ArcA modulates expression of virulence factors in Vibrio cholerae. Infect Immun 71: 5583-5589.
167. Lilley BN, Bassler BL (2000) Regulation of quorum sensing in Vibrio harveyi by LuxO and sigma-54. Molecular microbiology 36: 940-954.
168. Vance RE, Zhu J, Mekalanos JJ (2003) A constitutively active variant of the quorum-sensing regulator LuxO affects protease production and biofilm formation in Vibrio cholerae. Infection and immunity 71: 2571-2576.
169. Bassler BL, Wright M, Silverman MR (1994) Sequence and function of LuxO, a negative regulator of luminescence in Vibrio harveyi. Molecular microbiology 12: 403-412.
170. Vecerek B, Moll I, Afonyushkin T, Kaberdin V, Blasi U (2003) Interaction of the RNA chaperone Hfq with mRNAs: direct and indirect roles of Hfq in iron metabolism of Escherichia coli. Mol Microbiol 50: 897-909.
171. Ishikawa H, Otaka H, Maki K, Morita T, Aiba H (2012) The functional Hfq-binding module of bacterial sRNAs consists of a double or single hairpin preceded by a U-rich sequence and followed by a 3' poly(U) tail. RNA 18: 1062-1074.
172. Contag CH, Contag PR, Mullins JI, Spilman SD, Stevenson DK, et al. (1995) Photonic detection of bacterial pathogens in living hosts. Mol Microbiol 18: 593-603.
173. Wiles S, Pickard KM, Peng K, MacDonald TT, Frankel G (2006) In vivo bioluminescence imaging of the murine pathogen Citrobacter rodentium. Infect Immun 74: 5391-5396.
174. Levin JZ, Yassour M, Adiconis X, Nusbaum C, Thompson DA, et al. (2010) Comprehensive comparative analysis of strand-specific RNA sequencing methods. Nature methods 7: 709-715.

175. Mandlik A, Livny J, Robins WP, Ritchie JM, Mekalanos JJ, et al. (2011) RNA-Seq-based monitoring of infection-linked changes in Vibrio cholerae gene expression. Cell host & microbe 10: 165-174.
176. De Souza Silva O, Blokesch M (2010) Genetic manipulation of Vibrio cholerae by combining natural transformation with FLP recombination. Plasmid 64: 186-195.
177. Zuker M (2003) Mfold web server for nucleic acid folding and hybridization prediction. Nucleic acids research 31: 3406-3415.
178. Shao Y, Bassler BL (2012) Quorum-sensing non-coding small RNAs use unique pairing regions to differentially control mRNA targets. Molecular microbiology 83: 599-611.
179. Livny J, Brencic A, Lory S, Waldor MK (2006) Identification of 17 Pseudomonas aeruginosa sRNAs and prediction of sRNA-encoding genes in 10 diverse pathogens using the bioinformatic tool sRNAPredict2. Nucleic acids research 34: 3484-3493.
180. Said N, Rieder R, Hurwitz R, Deckert J, Urlaub H, et al. (2009) In vivo expression and purification of aptamer-tagged small RNA regulators. Nucleic Acids Res 37: e133.
181. Papenfort K, Podkaminski D, Hinton JC, Vogel J (2012) The ancestral SgrS RNA discriminates horizontally acquired Salmonella mRNAs through a single G-U wobble pair. Proceedings of the National Academy of Sciences of the United States of America 109: E757-764.
182. Queck SY, Jameson-Lee M, Villaruz AE, Bach TH, Khan BA, et al. (2008) RNAIII-independent target gene control by the agr quorum-sensing system: insight into the evolution of virulence regulation in Staphylococcus aureus. Molecular cell 32: 150-158.
183. Karaolis DK, Johnson JA, Bailey CC, Boedeker EC, Kaper JB, et al. (1998) A Vibrio cholerae pathogenicity island associated with epidemic and pandemic strains. Proceedings of the National Academy of Sciences of the United States of America 95: 3134-3139.
184. Guzman LM, Belin D, Carson MJ, Beckwith J (1995) Tight regulation, modulation, and high-level expression by vectors containing the arabinose PBAD promoter. Journal of bacteriology 177: 4121-4130.
185. Donnenberg MS, Kaper JB (1991) Construction of an eae deletion mutant of enteropathogenic Escherichia coli by using a positive-selection suicide vector. Infection and immunity 59: 4310-4317.
186. Terpe K (2003) Overview of tag protein fusions: from molecular and biochemical fundamentals to commercial systems. Appl Microbiol Biotechnol 60: 523-533.
187. Miller VL, Mekalanos JJ (1988) A novel suicide vector and its use in construction of insertion mutations: osmoregulation of outer membrane proteins and virulence determinants in Vibrio cholerae requires toxR. Journal of bacteriology 170: 2575-2583.
188. Pepke S, Wold B, Mortazavi A (2009) Computation for ChIP-seq and RNA-seq studies. Nat Methods 6: S22-32.
189. Camilli A, Mekalanos JJ (1995) Use of recombinase gene fusions to identify Vibrio cholerae genes induced during infection. Mol Microbiol 18: 671-683.

190. Trapnell C, Roberts A, Goff L, Pertea G, Kim D, et al. (2012) Differential gene and transcript expression analysis of RNA-seq experiments with TopHat and Cufflinks. Nature protocols 7: 562-578.
191. Morales VM, Backman A, Bagdasarian M (1991) A series of wide-host-range low-copy-number vectors that allow direct screening for recombinants. Gene 97: 39-47.