# Evolution of a *cis*-acting SNP that controls Type VI Secretion in *Vibrio cholerae*

Siu Lung Ng[1,3,4], Sophia Kammann[1,3,4], Gabi Steinbach[2,4], Tobias Hoffmann[1,3], Peter J Yunker[2,4], Brian K. Hammer[1,3,4#]

[1] School of Biological Sciences, Georgia Institute of Technology, Atlanta, Georgia, USA

[2]School of Physics, Georgia Institute of Technology, Atlanta, Georgia, USA

[3]Parker H. Petit Institute for Bioengineering & Bioscience, Georgia Institute of Technology, Atlanta, Georgia, USA

[4]Center for Microbial Diseases and Infection, Georgia Institute of Technology, Atlanta, Georgia, USA

#Corresponding author: brian.hammer@biology.gatech.edu

Running title: Evolution of the *V. cholerae* T6SS by a *cis*-acting SNP

**Abstract**

Mutations in regulatory mechanisms that control gene expression contribute to phenotypic diversity and thus facilitate the adaptation of microbes and other organisms to new niches. Comparative genomics can be used to infer rewiring of regulatory architecture based on large effect mutations like loss or acquisition of transcription factors but may be insufficient to identify small changes in non-coding, intergenic DNA sequence of regulatory elements that drive phenotypic divergence. In human-derived *Vibrio cholerae,* the response to distinct chemical cues triggers production of multiple transcription factors that can regulate the Type VI Secretion System (T6), a broadly distributed weapon for interbacterial competition. However, to date, the signaling network remains poorly understood because no regulatory element has been identified for the major T6 locus. Here we identify a conserved *cis*-acting single nucleotide polymorphism (SNP) controlling T6 transcription and activity. Sequence alignment of the T6 regulatory region from diverse *V. cholerae* strains revealed conservation of the SNP that we rewired to interconvert *V. cholerae* T6 activity between chitin-inducible and constitutive states. This study supports a model of pathogen evolution through a non-coding *cis*-regulatory mutation and preexisting, active transcription factors that confers a different fitness advantage to tightly regulated strains inside a human host and unfettered strains adapted to environmental niches.

**Importance**

Organisms sense external cues with regulatory circuits that trigger the production of transcription factors, which bind specific DNA sequences at promoters ("*cis*" regulatory elements) to activate target genes. Mutations of transcription factors or their regulatory elements create phenotypic diversity, allowing exploitation of new niches. Waterborne pathogen *Vibrio cholerae* encodes the Type VI Secretion System "nanoweapon" to kill competitor cells when activated. Despite identification of several transcription factors, no regulatory element has been identified in the promoter of the major Type VI locus, to date. Combining phenotypic, genetic, and genomic analysis of diverse *V. cholerae* strains, we discovered a single nucleotide polymorphism in the Type VI promoter that switches its killing activity between a constitutive state beneficial outside hosts and an inducible state for constraint in a host. Our results support a role for non-coding DNA in adaptation of this pathogen.

**Introduction**

A central role in the dynamic, temporal control of gene expression is played by transcription factors (TFs), diffusible "*trans*" products that bind to molecular switches within DNA sequences termed "*cis*"-regulatory elements (CREs). In eukaryotes, which lack horizontal gene transfer (HGT), mutations in CREs that alter TF binding sites are major contributors to

phenotypic diversity (1-3). In bacteria, pervasive HGT of TFs can alter entire regulatory circuits that allow adaptation to new niches, as prominently demonstrated in *Vibrio fischeri*, where host range is altered by the presence or absence of RcsS, a TF of biofilm and colonization genes (4, 5). By contrast, specific mutations at CREs in non-coding DNA are more difficult to identify and receive less attention as drivers of phenotypic divergence and evolutionary adaptation (6). Thus, elucidation of how microbes adapt to new niches, a process of fundamental importance in bacterial pathogenesis, requires coupling of genome-wide computational methods with experimental approaches to map the *cis*- and *trans*-regulatory interactions across and within species.

To understand how mutations play a role in microbial adaptation, pathogenic viruses and bacteria with lifestyles that exploit niches within and outside a human host are of great interest. Following ingestion, pandemic strains of the bacterium *Vibrio cholerae* can colonize the human gastrointestinal tract and secrete the cholera toxin that leads to the often fatal diarrhea responsible for seven pandemics to date (7-9). Conversely, *V. cholerae* isolated from non-human niches lack the horizontally-acquired prophage that carries the cholera toxin, and cause mild illness (10). By contrast, all sequenced *V. cholerae* encode a Type VI Secretion System (T6), a broadly distributed “nano-harpoon” weapon that injects toxic effector proteins into neighboring bacterial cells, leading to cell envelope damage and cell lysis (11, 12). Due to

its broad distribution among bacteria including those of the human gut, there is intense interest in understanding the T6 interactions between our microbiota and foreign pathogens, and whether they can be manipulated to influence health (13).

*V. cholerae* obtained from humans carry a limited arsenal of effectors and a T6 believed to be tailored for *in vivo* success (11, 14-19), while strains from non-human niches encode a more diverse effector repertoire (11, 14, 20, 21). To date, however, adaptative evolution mechanisms of T6 regulation in *V. cholerae* derived from non-human sources have largely been overlooked. Studies of human-derived strains identify two primary TFs for T6 activation (22-26). T6 control in pandemic strains (e.g. C6706 and A1552) requires QstR (Quorum-Sensing and Chitin-Dependent Regulator), which integrates multiple external cues (27-30), and contains a DNA binding domain postulated to interact with a presumptive CRE of the major T6 gene cluster (23, 27). T6 regulation in non-pandemic strain V52, which causes mild disease, requires TfoY, modulatable by intracellular signals, including cyclic di-GMP (25, 26). How QstR and TfoY control T6 transcription remains elusive, with no T6 CRE yet described. Elucidation of the differences in intraspecies T6 regulatory mechanisms between diverse *V. cholerae* isolates will provide insights into how pathogens emerge from nonpathogenic progenitors.

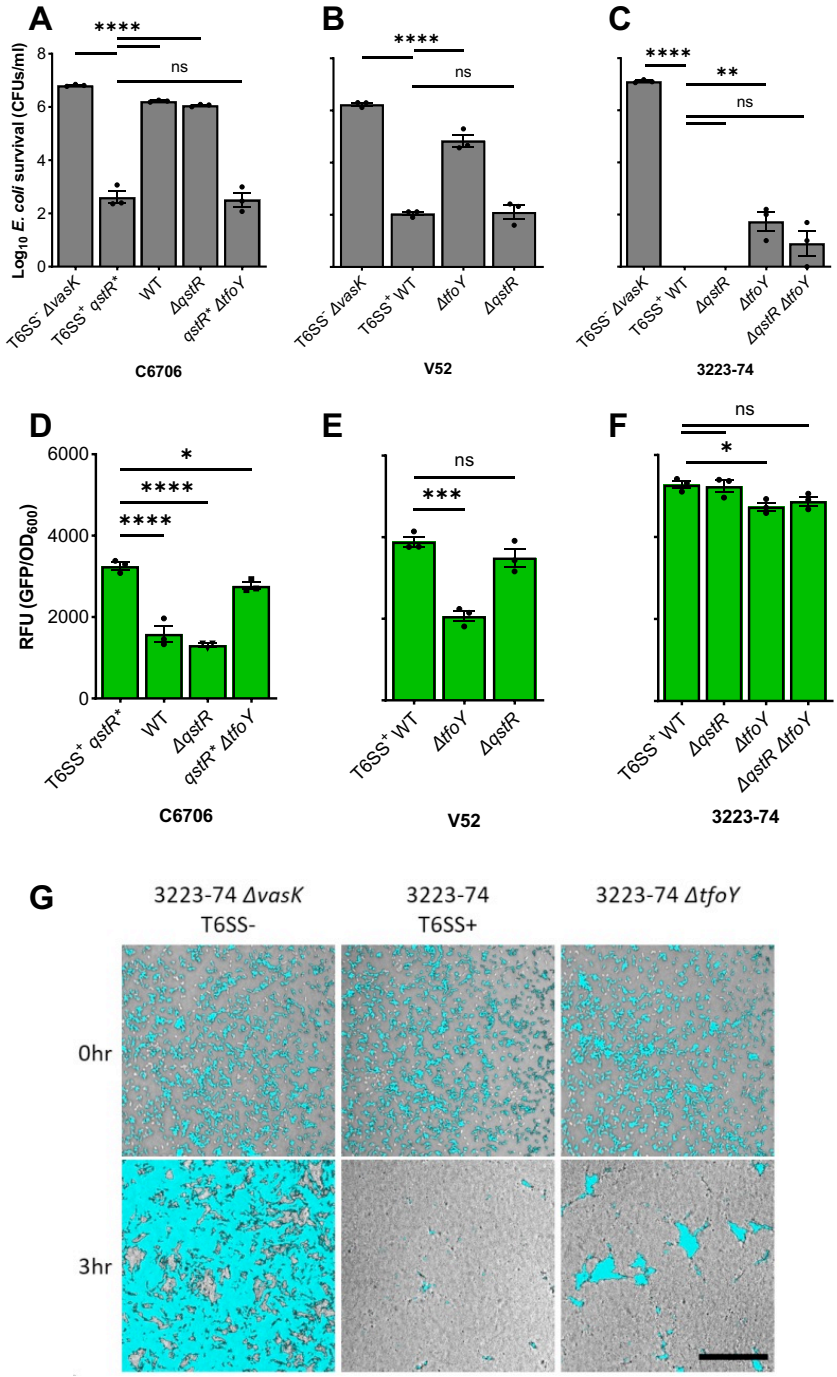
To understand the regulatory differences in *V. cholerae* strains, we examine here several environmental isolates that exhibit T6-mediated killing (31). Despite encoding functional signaling circuity and TFs, we find that QstR is dispensable for killing and that TfoY plays only a minor role in the strains tested. Thus, existing regulatory models fail to explain the T6 control in *V. cholerae* from human and non-human sources. Genomic analysis identifies one conserved non-coding single-nucleotide polymorphism (SNP) that we show interconverts *V. cholerae* T6 activity between chitin-inducible and constitutive states, which are QstR-dependent and TfoY-independent, respectively. We demonstrate that non-coding SNPs can rewire *cis*-regulatory elements to aid in adaptation of bacteria to different niches, including the human host.

## Results and Discussion

### Constitutive, *in vitro* T6 activity requires neither QstR nor TfoY.

In pandemic C6706, high cell density conditions (HCD) and chitin are required for induction of *qstR* which leads to activation of T6 genes. In the absence of chitin, C6706 with *qstR* expressed from a heterologous promoter (defined here as *qstR**) reduces survival of *Escherichia coli* "target" cells in co-culture by over 4-orders of magnitude (~10,000), compared to wildtype (WT) C6706, a T6$^-$ strain with a mutation in an essential structural gene (*ΔvasK*), and a strain with a *ΔqstR* mutation (Fig. 1A) (29). Deletion of *tfoY* does not reduce the killing activity of the

109 T6$^{+}$ *qstR** strain, but eliminates the robust killing in the non-pandemic strain V52 (serogroup
110 O37), which requires TfoY but not QstR (Fig. 1B) (26).
111
112 To determine whether QstR or TfoY participates in control of the T6 in non-human strains, we
113 examined 3223-74, a genetically-amenable, T6-proficient environmental strain (31). Like V52,
114 3223-74 does not require QstR to efficiently kill *E. coli* in conditions without chitin, but
115 surprisingly, also does not require TfoY. Isogenic strains carrying the *ΔtfoY* and *ΔqstR ΔtfoY*
116 mutations retain >99.99% killing activity, with only modest *E. coli* survival (Fig. 1C). Gene
117 fusions of the 5' intergenic region (IGR) of the major T6 cluster of each strain fused to green
118 fluorescent protein (*gfp*) confirm that transcriptional differences account for the killing
119 observed, with maximal *gfp* expression mirroring activity (i.e. low *E. coli* survival with high *gfp*
120 expression, and *vice versa*) (Fig. 1D-F). Confocal microscopy reinforces the negligible role of
121 TfoY on killing by 3223-74, with a Δ*tfoY* mutation having little effect on killing WT (Fig. 1G).
122 Transcription of plasmid-borne reporters is significantly higher in *V. cholerae* than in *E. coli*
123 (Fig. S1), supporting a hypothesis that an additional *V. cholerae*-specific regulator of the T6
124 may remain to be identified (Fig. S1).




125

126 **Figure 1. *Vibrio cholerae* 3223-74 T6 activity is QstR- and TfoY-independent.** (A-C) *V. cholerae*

127 strains with the indicated genotypes were co-cultured with chloramphenicol resistant ($Cm^r$) *E.*

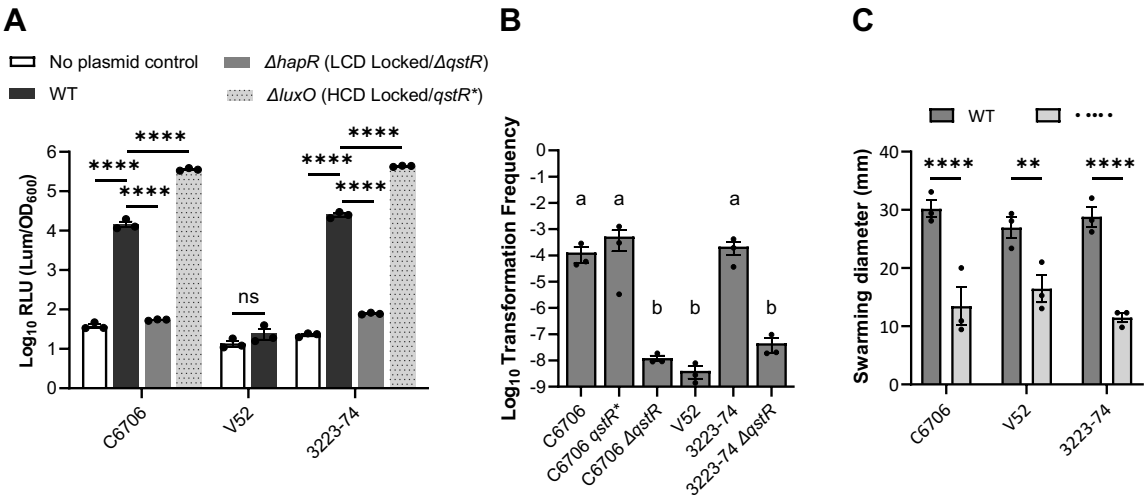
128 *coli* followed by determination of *E. coli* survival by counting of colony forming units (CFUs) on

129 LB agar with Cm. (D-F) Fluorescence levels are from reporters with *gfp* fused to the intergenic

region 5' of *vipA* derived from the three strains shown. The mean value ± S.E. of three independent co-cultures (A-C) and monocultures (D-F) are shown from one experiment, with similar results obtained in at least two other independent experiments. A one-way ANOVA with Dunnett post-hoc test was conducted to determine the significance: ns denotes not significant, ****p ≤ 0.0001, ***p ≤ 0.001, **p ≤ 0.01, *p ≤ 0.05. (G) *E. coli* cells expressing constitutive *gfp* were competed against 3223-74, with the same frame imaged at 0 h and 3 h by confocal microscopy. In the images, *gfp* signal from the *E. coli* is overlaid on top of bright-light images of the co-culture. Scale bar = 50 µm.

To probe each strain's T6-related regulatory circuitry, we measured canonical behaviors under control of HapR, QstR and TfoY; quorum sensing (QS) controlled bioluminescence, natural transformation, and motility, respectively (32-34). As expected, each TF is intact in C6706; but like several *V. cholerae* strains, V52 lacks a functional *hapR* gene that prevents QS and natural transformation (35, 36). Nonetheless, V52 encodes a functional *tfoY* that controls motility (Fig. 2A-B) (37). Interestingly, the regulatory circuity of *V. cholerae* 3223-74 is intact, like C6706, confirming that it encodes functional TFs (Fig. 2C), which are nonetheless expendable for T6-mediated killing. Because transposon mutagenesis failed to identify a novel T6 activator (not shown), we suspect regulation may be complex, perhaps involving more than one TF specific to *V. cholerae*. Nucleoid Associated Proteins (NAPs) that bind DNA both specifically and non-

149 specifically (38) may also contribute, since they are present in both species, likely regulated
150 differently (39), and participate in regulation of many promoters in numerous bacteria
151 including *Vibrios* (40).



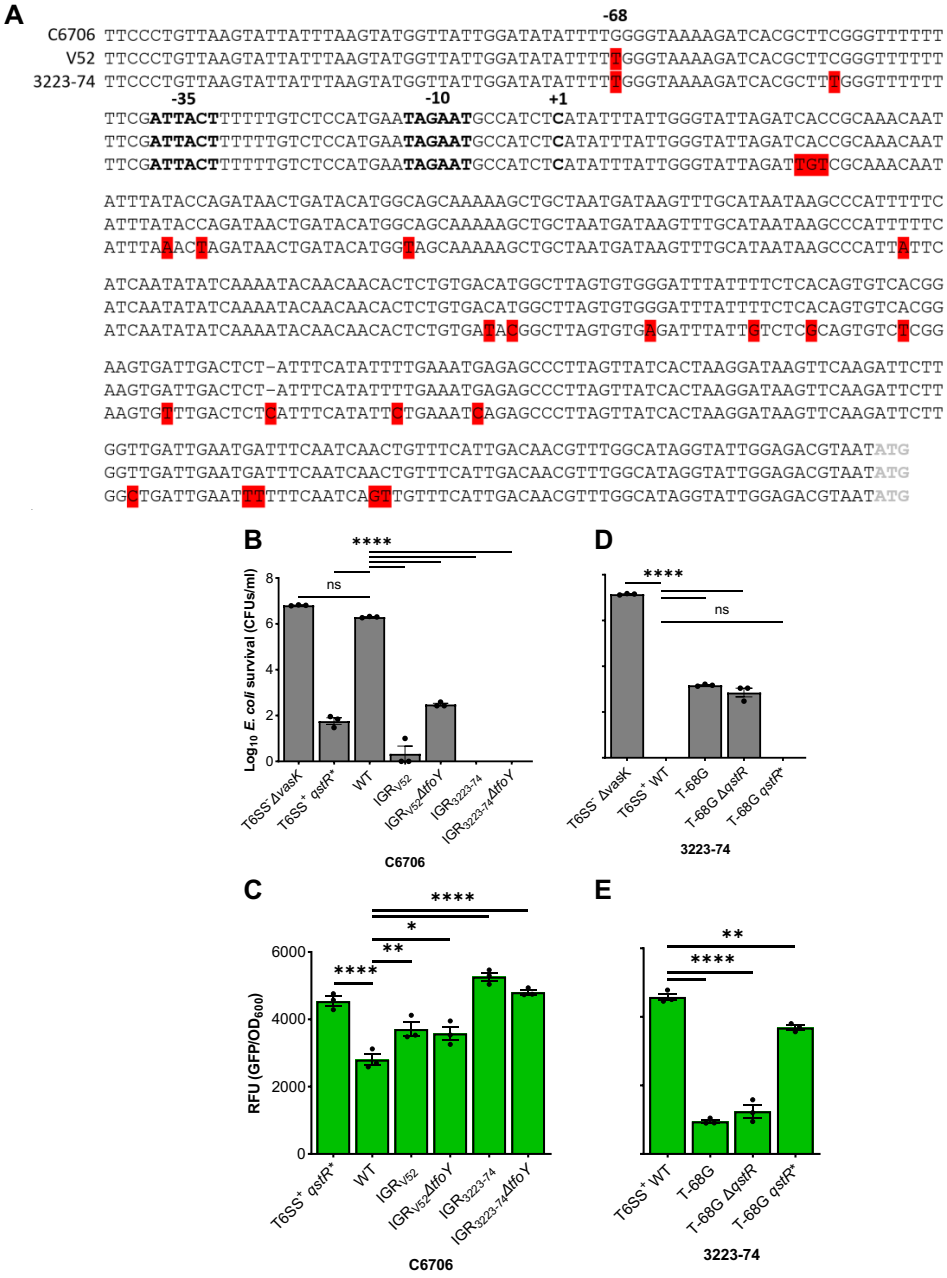

152

153 **Figure 2. *Vibrio cholerae* 3223-74 encodes functional HapR, QstR, and TfoY.** (A) *V. cholerae*
154 strains were grown in liquid LB with relative luminescence units per $OD_{600}$ measured at HCD
155 ($OD_{600}$ = 0.6-0.8). Statistical analyses were conducted with one-way ANOVA with Tukey post-
156 hoc test (C6706 and 3223-74) and one-tailed student's t-test (V52). LCD – Low Cell Density. (B)
157 *V. cholerae* strains with the indicated genotypes were grown in ASW with crab shell and
158 exogenous Spec-marked genomic DNA. Transformation frequency = $Spec^r$ CFU $ml^{-1}$ / total CFU
159 $ml^{-1}$. Statistical analyses were conducted with one-way ANOVA with Tukey post-hoc test.
160 Significance is denoted by letters. (C) *V. cholerae* strains were inoculated on 0.3% LB agar and
161 grew overnight. Statistical analyses were conducted with one-tailed student's t-test. Colony
162 diameters were physically measured from the furthest edges. All data shown are the mean ±

S.E. from one experiment, with similar results were obtained in at least two other independent experiments. ns: not significant, ****p ≤ 0.0001, **p ≤ 0.01.

A SNP in the T6 intergenic region confers QstR-dependency.

Human and environmental isolates of *V. cholerae* we have characterized prior (31) share ≥97% average nucleotide identity with many chromosomal differences (11), but inspection of the T6 IGRs of C6706, V52 and 3223-74 revealed only 17 SNPs and 3 multinucleotide polymorphisms (Fig. 3A), which we hypothesized could contribute to the differences in T6 transcription and killing activity observed. To address this, we replaced the T6 IGR of C6706 on the chromosome with that from V52 and 3223-74 and measured killing activity. While C6706 carrying the *qstR** allele, but not WT, adeptly kills *E. coli*, both IGR replacements increase the killing efficiency of WT C6706 by 5- to 6-orders of magnitude (Fig. 3B), mimicking the robust killing observed by WT V52 and 3223-74 (Fig. 1B-C). Deletion of *tfoY* in C6706 with V52's IGR increases *E. coli* survival (~ 2-logs), as observed with V52, but does not alter *E. coli* survival with 3223-74's IGR (Fig. 3B). Chromosomal transcriptional *gfp* reporters with identical mutations were elevated relative to WT C6706 in each IGR replacement strain (Fig. 3C), consistent with the enhanced killing detected. These results support a hypothesis that a novel CRE lies within the IGR 5' of the T6 locus, despite a lack of any known direct TF-DNA interactions at this locus identified to date.




182

183 **Figure 3. G-388T mutation abolishes QstR dependence in C6706 and T-388G confers QstR**

184 **dependence to 3223-74.** (A) Alignment of the IGR upstream of *vipA* was conducted using

185 MUSCLE. SNPs and MNPs are highlighted in red, one gap indicated with a "–", the putative

186 promoter and the transcriptional start site (TSS; +1) in bold, and the start codon of *vipA* in

187 grey. (B) the C6706 5' IGR of *vipA* was replaced with the IGR from either V52 or 3223-74. (D)

188 A T-68G mutation in the 5' IGR of *vipA* was introduced into 3223-74 with different *qstR* alleles.
189 Competition assays were conducted by co-culturing *V. cholerae* killers and Cm$^r$ *E. coli* target
190 followed by determination of *E. coli* survival by counting of colony forming units (CFUs) on LB
191 agar with Cm. (C, E) Shown are fluorescence levels of transcriptional reporters with *gfp* fused
192 to corresponding IGRs of *vipA* expressed in either C6706 (C) or 3223-74 (E). The mean value ±
193 S.E. of three independent co-cultures (B and D) and monocultures (C and E) are shown from
194 one experiment, with similar results were obtained in at least two other independent
195 experiments. A one-way ANOVA with Dunnett post-hoc test was conducted to determine the
196 significance - ns: not significant, ****p ≤ 0.0001, **p ≤ 0.01, *p ≤ 0.05.
197
198 To begin mapping the T6 IGR region and SNP locations, we experimentally determined the
199 transcriptional start site (+1) by 5' Rapid Amplification of cDNA Ends (Methods). The +1 of
200 transcription resides 320 nucleotides (nt) 5' of the ATG of the first T6 gene (*vipA*), and adjacent
201 to a putative promoter with 8/12 identical nts compared to the consensus sigma70-
202 dependent promoter (Fig. 3A). The +1 is consistent with paired-end RNAseq results we have
203 reported prior (29). Because the majority of 5' untranslated regions (UTRs) in *V. cholerae* are
204 20-40 nt, with few exceeding 300 nt (41), we speculate that the 320 nt 5' UTR of the major T6
205 gene cluster may be post-transcriptionally regulated, beyond the sRNA interactions already
206 described near the ribosome binding site (RBS) (42). Alignment of the IGRs of C6706 and V52

reveals a single SNP at -68, with a guanine (G) in C6706 at that position and a thymine (T) in V52 (Fig. 3A).

The replacement of the C6706 IGR with V52 was effectively a G-68T mutation (Fig. 3B-C), thus we further tested whether G was necessary for QstR activation by replacing the T with a G at position -68 (T-68G) in the 3223-74 WT, *qstR**, and *ΔqstR* backgrounds. The T-68G mutation significantly increases *E. coli* survival and decreases T6 transcription in WT 3223-74 and the *ΔqstR* derivative, with killing restored in the strain with the *qstR** allele (Fig. 3D-E). Thus, a G at position -68 confers inducible, QstR-control, while a T results in constitutive killing *in vitro*, consistent with results recently reported (43). Based on these results we predicted this SNP is a result of adaptive evolution to control T6 activity in different environments.

The SNP at -68 is evolutionarily conserved.

To determine whether the SNP at -68 is prevalence in *V. cholerae*, we aligned the T6 IGR sequences of diverse strains that we have characterized prior for T6 killing activity (Fig. 4A) (31). Consistent with prior studies (11, 14, 16, 18), our phylogenetic analysis (Methods) of the T6 IGRs places human strains in a distinct clade, with the exception of two O1 strains isolated nearly a century ago (NCTC8457 and MAK757), and two non-O1 strains (MZO-2 O14 and V52 O37; Fig. S2). All 23 environmental isolates carry the T-68 SNP and displays constitutive T6

226 activity, with one exception that is chitin-inducible (1496-86) (Fig. 4A, S4). By contrast, the 18
227 human isolates tested carry either G or T at the -68 position (Fig. 4A, S5). The 13 chitin-
228 inducible human isolates carry a G; five show constitutive activity and carry a T like
229 environmental strains, with one exception that is constitutive yet carries the G (2010EL-1749)
230 (Fig. 4A, S5). Neither C nor A are observed at -68 in any stains tested, although both pyrimidine
231 nucleotides (T and C) confer constitutive killing at -68, and both purines (G and A) behave
232 similarly (Fig. S3). The focal SNP location is distal from the promoter, but inconsistent with AT-
233 rich "UP elements" that reside immediately upstream of the promoter at -38 to -59 and
234 interact directly with the alpha subunit of RNAP (44). We propose the SNP is more likely a
235 component of a CRE for a TF to be determined. Indeed, transversion mutations have greater
236 effects of TF binding than transitions, as noted here (Fig. S3) likely due to changes in shape of
237 the DNA backbone or DNA-amino acid contacts (45, 46).
238
239 We examined regulation of three additional genetically manipulatable environmental strains
240 (VC22, 2479-89, and 2512-86) that exhibit T6 killing (31). Like 3223-74, QstR is expendable in
241 each strain (Fig. 4B-E) while TfoY contributes to some extent in activating T6, with varying *E.*
242 *coli* recovery observed in each derivative carrying the *ΔtfoY* mutation (Fig. 4B-E). Taken
243 together, our findings reveal that the constitutive T6 killing activity of environmental *V.*
244 *cholerae* is driven by a T at position -68, which obviates the QstR requirement, and permits

modest TfoY regulation.

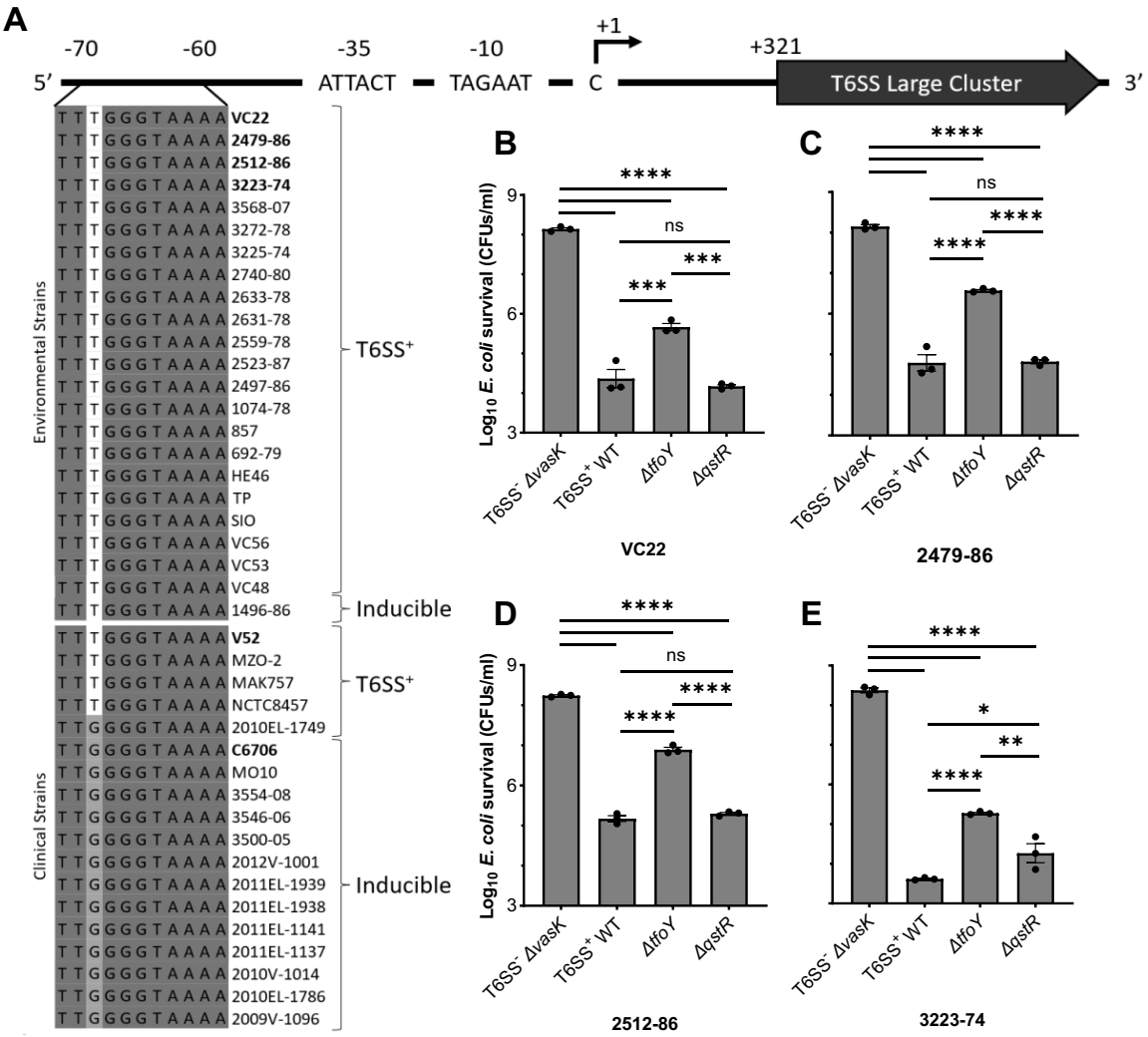



**Figure 4. Environmental *V. cholerae* isolates encode a T at position -68 while human, chitin-induced isolates encode a G.** (A) A SNP at position -68 in the IGR of the major T6 cluster controls killing activity. Conserved nts are in dark grey and the SNP of interest is highlighted in white/grey. T6 control was categorized as described (31). (B-E) Survival of *E. coli* following competition assays with WT *V. cholerae* strains and mutants was determined by CFU counts. Data shown are the mean ± S.E. of three independent experiments. A one-way ANOVA with

253 Tukey post-hoc test was conducted to determine the significance - ns: not significant, ****p ≤
254 0.0001, ***p ≤ 0.001, **p ≤ 0.01, *p ≤ 0.05.
255
256 Bacterial adaptation to unexploited niches can be the result of horizontal gene transfer events
257 (5) as well as mutations in protein coding and promoter regions (47, 48). Here we describe an
258 intergenic non-coding SNP that coordinates adaptation by altering T6 control between two
259 states – one that in inducible and the other that displays constitutive activity. While the first
260 Type VI Secretion System was first described in *V. cholerae* in 2006, the knowledge of its
261 regulation remains largely restricted to human isolates and incomplete, with the identity of a
262 TF that directly controls the major T6 cluster elusive to this date (22, 24). We speculate that
263 the focal SNP we identified at position -68 is a component of a CRE that contributes to
264 pathoadaptation (Fig. 3A), a result of adaptive evolution, which allows *V. cholerae* to carefully
265 control the T6SS expression in specific environments. Our results are consistent with the
266 hypothesis that constitutive T6SS is beneficial in aquatic environments outside a human host
267 (49), with varying degrees of TfoY contribution, which may act directly or indirectly at the
268 transcriptional or posttranscriptional level (Fig. 3A and Fig. 4B-E, S4, S5). During human
269 infection where selection promotes dampened T6SS, *V. cholerae* with a T-to-G mutation
270 (inducible T6) are favored. In fact, T6-deficient human isolates (e.g. O395) have been reported
271 to have less competitive fitness in human intestinal colonization and infection (19, 50).

Although low level, basal expression of T6 contributes to pathogenesis of C6706 (51), overexpression of T6SS may be deleterious in vivo. Indeed, we have reported prior that *V. cholerae* with constitutive T6SS induces violent peristaltic contractions in a fish host (52), which may disrupt the interaction between *V. cholerae* and the gut microflora.

There remains a pressing public health need to understand the emergence of pathogens from environmental reservoirs (53). Efforts such as Microbial Genome Wide Association Studies (54) to identify genetic variants in genomes that are associated with phenotypes like virulence and antibiotic sensitivity, will be bolstered by knowledge of the ecological and evolutionary processes that promote pathogen-host association. Defining the plasticity of the regulatory circuity controlling the T6 weapon will provide insights into the role of polymorphisms in the evolution of this and other pathogens.

## Materials and Methods

### Bacterial growth conditions and plasmid constructions

All *V. cholerae* and *E. coli* (Table S1) strains were grown aerobically at 37 °C overnight in Lysogeny Broth (LB) with constant shaking or statically on LB agar. Ampicillin (100 µg/ml), kanamycin (50 µg/ml), chloramphenicol (10 µg/ml), spectinomycin (100 µg/ml), streptomycin (5 mg/ml), sucrose (20% w/v) and diaminopimelic acid (50 µg/ml) were supplemented where

appropriate.

Plasmids (Table S2) used were constructed with DNA restriction nucleases (Promega – WI, USA), Gibson Assembly mix (New England Biolabs – MA, USA), and PCR amplification (Qiagen - Hilden, Germany) by PCR with Q5 polymerase (New England Biolabs – MA, USA), and primers (Table S3) generated by Eton Bioscience Inc (NC, USA) or Eurofins Genomics (KY, USA). All reagents were used according to the manufacturer's instructions. Plasmids were confirmed by PCR and Sanger sequencing by Eton Bioscience Inc (NC, USA).

*V. cholerae* mutant construction

All genetically engineered strains of *V. cholerae* were constructed with established allelic exchange methods using vector pKAS32 (55) and pRE118 (Addgene - Plasmid #43830). All Insertions, deletions, and mutations were confirmed by PCR and Sanger sequencing conducted by Eton Bioscience Inc (NC, USA). Primers used are in Table S3.

Fluorescence microscopy

*V. cholerae* 3223-74 strains and chromosomal-labeled GFP *E. coli* were separately back-diluted 1:100 and incubated at 37 °C for 3 h. *V. cholerae* and *E. coli* were normalized to $OD_{600}$ = 1 and mixed in a 1:5 ratio. A 2 µL aliquot of a mixed culture was inoculated on LB agar and allowed

to dry. Cells were imaged before and after a 3 h incubation at 37 °C and 96-100% humidity using an Eclipse Ti-E Nikon (NY, USA) inverted microscope with a Perfect Focus System and camera previously described (11). The images were processed with ImageJ (34).

Motility assay

Overnight cultures of *V. cholerae* were diluted to $OD_{600}$ = 0.1, and 1 µL inoculated onto pre-dried LB plates with 0.3 % agar. Cells were incubated at 37 °C statically overnight, with motility determined by measuring the swarming diameter.

Transformation assay

Chitin-induced transformation frequency was measured as described with defined artificial sea water (450 mM NaCl, 10 mM KCl, 9 mM $CaCl_2$, 30 mM $MgCl_2{\cdot}6H_2O$, and 16 mM $MgSO_4{\cdot}7H_2O$; pH 7.8) (56). Bacteria were incubated with extracellular DNA in triplicate wells containing crab shell tabs, and transformation frequency calculated as Spectinomycin resistant ($Spec^r$) CFU $ml^{-1}$ / total CFU $ml^{-1}$.

QS-dependent Luciferase assay

Overnight cultures of the bacterial strains were diluted to $OD_{600}$ = 0.001 in liquid LB in microtiter plates and incubated at 37 °C with shaking. The $OD_{600}$ and luminescence were

329 measured each h with a BioTek (VT, USA) Synergy H1 microplate reader to calculate Relative
330 Luminescence Units (RLU) as Luminescence/$OD_{600}$. Data were collected when $OD_{600}$ = 0.6-0.8.
331 LB medium was used as the blank for the $OD_{600}$ and luminescence.
332
333 GFP transcriptional reporter quantification
334 Overnight cultures of *V. cholerae* or *E. coli* were diluted 1:100 and incubated at 37 °C for 3 h.
335 To enhance the translation of *gfp*, the sequence of the native RBS (12 nt sequence) was
336 replaced with the T7 RBS (12 nt sequence) in the primers used to make the fusions. Cm was
337 added to maintain the plasmid-borne versions of reporters that were cloned into plasmid
338 pSLS3. 300 µL aliquots were transferred to black microtiter plates to read the $OD_{600}$ and GFP
339 fluorescence (Excitation: 485, Emission: 528) with a BioTek Synergy H1 microplate reader (VT,
340 USA) to calculate Relative Luminescence Units (RLU) as Luminescence/$OD_{600}$. LB medium was
341 used as the blank for the $OD_{600}$. Strain lacking reporters served as blanks for GFP fluorescence.
342 RFU was calculated by blanked GFP fluorescence / blanked $OD_{600}$.
343
344 T6-mediated killing assay
345 Overnight cultures of *V. cholerae* or *E. coli* were back-diluted 1:100 and incubated at 37 °C for
346 3 h. *V. cholerae* strains and the $Cm^r$ *E. coli* target were normalized to $OD_{600}$ = 1 and then mixed
347 at a ratio of either 10:1 or 1:5. A 50 µL mixed culture was spotted onto LB agar and dried. After

a 3 h incubation at 37°C, cells were resuspended in 5 ml of LB, and serial dilutions were conducted. Finally, the resuspension was inoculated on a LB agar containing Cm to select for the surviving *E. coli*, which was incubated overnight at 37 °C and the *E. coli* colonies were counted and shown as CFU $mL^{-1}$.

RNA extraction and determination of the +1 of transcription by 5'-RACE

Overnight cultures of *V. cholerae* were back-diluted 1:100 and incubated at 37 °C for 3 h before lysing. Three independent cultures of T6-active *V. cholerae* C6706 *qstR** and 3223-74 WT were harvested by centrifugation at room temperature. RNA isolation, genomic DNA removal, and RNA clean-up were performed as previously described (57). Genomic DNA contamination was confirmed by conducting PCR with primer pair specific for 16S rRNA loci (*rrsA*) as previously described (Table S3) (58). RNA purity was confirmed by NanoDrop (260 / 280 ≈ 2.0).

5'-RACE (Invitrogen™ - MA, USA) was conducted according to the manufacturer's protocol with slight modifications. Specifically, SuperScript™ IV reverse transcriptase (Invitrogen™ - MA, USA) was used to complete the first strand cDNA synthesis. Two *vipA*-specific primers (GT3056 and GT3060) were used to identify the +1 of transcription for the major T6 gene cluster (Table S3). PCR products were purified with QIAquick PCR purification kit (Qiagen - Hilden, Germany) or Zymoclean gel DNA recovery kit (Zymo Research - CA, USA). Sanger

sequencing was conducted by Eton Bioscience Inc. (NC, USA) with the corresponding nesting primer (Table S3).

Genomic and phylogenetic analysis

Genome sequences of *V. cholerae* strains were collected from NCBI Genome database (Table S4) (59). The IGR upstream of major T6 cluster was extracted, aligned, and presented using BLAST+ v2.2.18 (60), MUSCLE v3.8 (https://www.ebi.ac.uk/Tools/msa/muscle/) (61, 62), and ESPript 3.0 (https://espript.ibcp.fr/) (37). The DNA sequence of the IGR was used for phylogenetic analysis, and the phylogenetic tree was constructed by the Maximum likelihood method using MEGA11 (63, 64).

For 2012V-1001, 2011EL-1939, 2011EL-1938, and 2011EL-1141 that do not have genome sequence available, colony PCR was conducted to amplify the 5' IGR of the major T6 cluster using OneTaq DNA Polymerase (New England Biolabs – MA, USA). PCR products were confirmed with gel electrophoresis and Sanger sequencing by Eton Bioscience Inc. (NC, USA) with the identical primer pair (Table S3).

**Acknowledgements**

We would like to thank Dr. Jyl S. Matson for assistance with RNA isolation and Dr. Marvin Whiteley and current and past members of the Hammer Lab for critiques and discussion, specifically, Dr. Samit Watve and Rakin Choudhury for bioinformatic advice and assistance.

**Competing interests**

The authors have no competing interests.

391 **References**

392 1. Wray GA. The evolutionary significance of cis-regulatory mutations. Nature reviews
393 Genetics. 2007;8(3):206-16. Epub 2007/02/17. doi: 10.1038/nrg2063. PubMed PMID:
394 17304246.

395 2. Stern DL, Orgogozo V. THE LOCI OF EVOLUTION: HOW PREDICTABLE IS GENETIC
396 EVOLUTION? Evolution. 2008;62(9):2155-77. doi: https://doi.org/10.1111/j.1558-
397 5646.2008.00450.x.

398 3. Wittkopp PJ, Kalay G. Cis-regulatory elements: molecular mechanisms and evolutionary
399 processes underlying divergence. Nature reviews Genetics. 2011;13(1):59-69. Epub
400 2011/12/07. doi: 10.1038/nrg3095. PubMed PMID: 22143240.

401 4. Norsworthy AN, Visick KL. Signaling between two interacting sensor kinases promotes
402 biofilms and colonization by a bacterial symbiont. Mol Microbiol. 2015;96(2):233-48. Epub
403 2015/01/15. doi: 10.1111/mmi.12932. PubMed PMID: 25586643; PMCID: PMC4465548.

404 5. Mandel MJ, Wollenberg MS, Stabb EV, Visick KL, Ruby EG. A single regulatory gene is
405 sufficient to alter bacterial host range. Nature. 2009;458(7235):215-8. Epub 2009/02/03. doi:
406 10.1038/nature07660. PubMed PMID: 19182778; PMCID: PMC2713604.

407 6. Perez JC, Groisman EA. Evolution of transcriptional regulatory circuits in bacteria. Cell.
408 2009;138(2):233-44. Epub 2009/07/28. doi: 10.1016/j.cell.2009.07.002. PubMed PMID:
409 19632175; PMCID: PMC2726713.

410 7. Dziejman M, Balon E, Boyd D, Fraser CM, Heidelberg JF, Mekalanos JJ. Comparative
411 genomic analysis of Vibrio cholerae: genes that correlate with cholera endemic and pandemic
412 disease. Proceedings of the National Academy of Sciences of the United States of America.
413 2002;99(3):1556-61. Epub 2002/01/31. doi: 10.1073/pnas.042667999. PubMed PMID:
414 11818571; PMCID: PMC122229.
415 8. Faruque SM, Albert MJ, Mekalanos JJ. Epidemiology, Genetics, and Ecology of
416 ToxigenicVibrio cholerae. Microbiology and Molecular Biology Reviews.
417 1998;62(4):1301-14. doi: doi:10.1128/MMBR.62.4.1301-1314.1998.
418 9. Karaolis DKR, Johnson JA, Bailey CC, Boedeker EC, Kaper JB, Reeves PR. A Vibrio
419 cholerae pathogenicity island associated with epidemic and pandemic strains.
420 Proceedings of the National Academy of Sciences. 1998;95(6):3134-9. doi:
421 10.1073/pnas.95.6.3134.
422 10. Pruzzo C, Vezzulli L, Colwell RR. Global impact of Vibrio cholerae interactions with chitin.
423 Environ Microbiol. 2008;10(6):1400-10. Epub 2008/03/04. doi: 10.1111/j.1462-
424 2920.2007.01559.x. PubMed PMID: 18312392.
425 11. Crisan CV, Chande AT, Williams K, Raghuram V, Rishishwar L, Steinbach G, Watve SS,
426 Yunker P, Jordan IK, Hammer BK. Analysis of Vibrio cholerae genomes identifies new type VI
427 secretion system gene clusters. Genome Biol. 2019;20(1):163. Epub 2019/08/14. doi:
428 10.1186/s13059-019-1765-5. PubMed PMID: 31405375; PMCID: PMC6691524.

429 12. MacIntyre DL, Miyata ST, Kitaoka M, Pukatzki S. The Vibrio cholerae type VI secretion
430 system displays antimicrobial properties. Proceedings of the National Academy of Sciences of
431 the United States of America. 2010;107(45):19520-4. Epub 2010/10/27. doi:
432 10.1073/pnas.1012931107. PubMed PMID: 20974937; PMCID: PMC2984155.
433 13. Gallegos-Monterrosa R, Coulthurst SJ. The ecological impact of a bacterial weapon:
434 microbial interactions and the Type VI secretion system. FEMS Microbiol Rev. 2021;45(6). Epub
435 2021/06/23. doi: 10.1093/femsre/fuab033. PubMed PMID: 34156081; PMCID: PMC8632748.
436 14. Hussain NAS, Kirchberger PC, Case RJ, Boucher YF. Modular Molecular Weaponry Plays a
437 Key Role in Competition Within an Environmental Vibrio cholerae Population. Frontiers in
438 Microbiology. 2021;12(979). doi: 10.3389/fmicb.2021.671092.
439 15. Ishikawa T, Sabharwal D, Bröms J, Milton DL, Sjöstedt A, Uhlin BE, Wai SN, Camilli A.
440 Pathoadaptive Conditional Regulation of the Type VI Secretion System in Vibrio cholerae O1
441 Strains. Infection and Immunity. 2012;80(2):575-84. doi: doi:10.1128/IAI.05510-11.
442 16. Kirchberger PC, Unterweger D, Provenzano D, Pukatzki S, Boucher Y. Sequential
443 displacement of Type VI Secretion System effector genes leads to evolution of diverse
444 immunity gene arrays in Vibrio cholerae. Scientific Reports. 2017;7(1):45133. doi:
445 10.1038/srep45133.
446 17. Santoriello FJ, Pukatzki S. When the pandemic opts for the lockdown: Secretion system
447 evolution in the cholera bacterium. Microb Cell. 2021;8(3):69-72. Epub 2021/03/09. doi:

448 10.15698/mic2021.03.744. PubMed PMID: 33681340; PMCID: PMC7919388.

449 18. Unterweger D, Miyata ST, Bachmann V, Brooks TM, Mullins T, Kostiuk B, Provenzano D,
450 Pukatzki S. The Vibrio cholerae type VI secretion system employs diverse effector modules for
451 intraspecific competition. Nature Communications. 2014;5(1):3549. doi:
452 10.1038/ncomms4549.

453 19. Kostiuk B, Santoriello FJ, Diaz-Satizabal L, Bisaro F, Lee K-J, Dhody AN, Provenzano D,
454 Unterweger D, Pukatzki S. Type VI secretion system mutations reduced competitive fitness of
455 classical Vibrio cholerae biotype. Nature Communications. 2021;12(1):6457. doi:
456 10.1038/s41467-021-26847-y.

457 20. Crisan CV, Chandrashekar H, Everly C, Steinbach G, Hill SE, Yunker PJ, Lieberman RR,
458 Hammer BK, Ellermeier CD. A New Contact Killing Toxin Permeabilizes Cells and Belongs to a
459 Broadly Distributed Protein Family. mSphere. 2021;6(4):e00318-21. doi:
460 doi:10.1128/mSphere.00318-21.

461 21. Drebes Dörr NC, Blokesch M. Interbacterial competition and anti-predatory behaviour of
462 environmental Vibrio cholerae strains. Environmental Microbiology. 2020;22(10):4485-504.
463 doi: https://doi.org/10.1111/1462-2920.15224.

464 22. Crisan CV, Hammer BK. The Vibrio cholerae type VI secretion system: toxins, regulators
465 and consequences. Environ Microbiol. 2020;22(10):4112-22. Epub 2020/03/07. doi:
466 10.1111/1462-2920.14976. PubMed PMID: 32133757.

467 23. Jaskolska M, Stutzmann S, Stoudmann C, Blokesch M. QstR-dependent regulation of
468 natural competence and type VI secretion in Vibrio cholerae. Nucleic Acids Res.
469 2018;46(20):10619-34. Epub 2018/08/14. doi: 10.1093/nar/gky717. PubMed PMID:
470 30102403; PMCID: PMC6237807.
471 24. Joshi A, Kostiuk B, Rogers A, Teschler J, Pukatzki S, Yildiz FH. Rules of Engagement: The
472 Type VI Secretion System in Vibrio cholerae. Trends Microbiol. 2017;25(4):267-79. Epub
473 2016/12/29. doi: 10.1016/j.tim.2016.12.003. PubMed PMID: 28027803; PMCID:
474 PMC5365375.
475 25. Joshi A, Mahmoud SA, Kim SK, Ogdahl JL, Lee VT, Chien P, Yildiz FH. c-di-GMP inhibits
476 LonA-dependent proteolysis of TfoY in Vibrio cholerae. PLoS Genet. 2020;16(6):e1008897.
477 Epub 2020/06/27. doi: 10.1371/journal.pgen.1008897. PubMed PMID: 32589664; PMCID:
478 PMC7371385.
479 26. Metzger LC, Stutzmann S, Scrignari T, Van der Henst C, Matthey N, Blokesch M.
480 Independent Regulation of Type VI Secretion in Vibrio cholerae by TfoX and TfoY. Cell Rep.
481 2016;15(5):951-8. Epub 2016/04/28. doi: 10.1016/j.celrep.2016.03.092. PubMed PMID:
482 27117415; PMCID: PMC4858559.
483 27. Lo Scrudato M, Blokesch M. A transcriptional regulator linking quorum sensing and chitin
484 induction to render Vibrio cholerae naturally transformable. Nucleic Acids Res.
485 2013;41(6):3644-58. Epub 2013/02/06. doi: 10.1093/nar/gkt041. PubMed PMID: 23382174;

PMCID: PMC3616704.

28. Meibom KL, Blokesch M, Dolganov NA, Wu CY, Schoolnik GK. Chitin induces natural competence in Vibrio cholerae. Science. 2005;310(5755):1824-7. Epub 2005/12/17. doi: 10.1126/science.1120096. PubMed PMID: 16357262.

29. Watve SS, Thomas J, Hammer BK. CytR Is a Global Positive Regulator of Competence, Type VI Secretion, and Chitinases in Vibrio cholerae. PLoS One. 2015;10(9):e0138834. Epub 2015/09/25. doi: 10.1371/journal.pone.0138834. PubMed PMID: 26401962; PMCID: PMC4581735.

30. Tsou AM, Cai T, Liu Z, Zhu J, Kulkarni RV. Regulatory targets of quorum sensing in Vibrio cholerae: evidence for two distinct HapR-binding motifs. Nucleic Acids Res. 2009;37(8):2747-56. Epub 2009/03/12. doi: 10.1093/nar/gkp121. PubMed PMID: 19276207; PMCID: PMC2677876.

31. Bernardy EE, Turnsek MA, Wilson SK, Tarr CL, Hammer BK, Liu S-J. Diversity of Clinical and Environmental Isolates of Vibrio cholerae in Natural Transformation and Contact-Dependent Bacterial Killing Indicative of Type VI Secretion System Activity. Applied and Environmental Microbiology. 2016;82(9):2833-42. doi: doi:10.1128/AEM.00351-16.

32. Jobling MG, Holmes RK. Characterization of hapR, a positive regulator of the Vibrio cholerae HA/protease gene hap, and its identification as a functional homologue of the Vibrio harveyi luxR gene. Molecular Microbiology. 1997;26(5):1023-34. doi:

https://doi.org/10.1046/j.1365-2958.1997.6402011.x.

33. Pursley BR, Maiden MM, Hsieh ML, Fernandez NL, Severin GB, Waters CM. Cyclic di-GMP Regulates TfoY in Vibrio cholerae To Control Motility by both Transcriptional and Posttranscriptional Mechanisms. J Bacteriol. 2018;200(7). Epub 2018/01/10. doi: 10.1128/JB.00578-17. PubMed PMID: 29311281; PMCID: PMC5847662.

34. Schneider CA, Rasband WS, Eliceiri KW. NIH Image to ImageJ: 25 years of image analysis. Nat Methods. 2012;9(7):671-5. Epub 2012/08/30. doi: 10.1038/nmeth.2089. PubMed PMID: 22930834; PMCID: PMC5554542.

35. Zheng J, Shin OS, Cameron DE, Mekalanos JJ. Quorum sensing and a global regulator TsrA control expression of type VI secretion and virulence in Vibrio cholerae. Proceedings of the National Academy of Sciences. 2010;107(49):21128-33. doi: 10.1073/pnas.1014998107.

36. Joelsson A, Liu Z, Zhu J. Genetic and phenotypic diversity of quorum-sensing systems in clinical and environmental isolates of Vibrio cholerae. Infect Immun. 2006;74(2):1141-7. Epub 2006/01/24. doi: 10.1128/IAI.74.2.1141-1147.2006. PubMed PMID: 16428762; PMCID: PMC1360356.

37. Robert X, Gouet P. Deciphering key features in protein structures with the new ENDscript server. Nucleic Acids Res. 2014;42(Web Server issue):W320-4. Epub 2014/04/23. doi: 10.1093/nar/gku316. PubMed PMID: 24753421; PMCID: PMC4086106.

524 38. Holowka J, Zakrzewska-Czerwinska J. Nucleoid Associated Proteins: The Small Organizers
525 That Help to Cope With Stress. Front Microbiol. 2020;11:590. Epub 2020/05/07. doi:
526 10.3389/fmicb.2020.00590. PubMed PMID: 32373086; PMCID: PMC7177045.
527 39. Ayala JC, Silva AJ, Benitez JA. H-NS: an overarching regulator of the Vibrio cholerae life
528 cycle. Res Microbiol. 2017;168(1):16-25. Epub 2016/08/06. doi:
529 10.1016/j.resmic.2016.07.007. PubMed PMID: 27492955; PMCID: PMC5241179.
530 40. Chaparian RR, Olney SG, Hustmyer CM, Rowe-Magnus DA, van Kessel JC. Integration host
531 factor and LuxR synergistically bind DNA to coactivate quorum-sensing genes in Vibrio harveyi.
532 Mol Microbiol. 2016;101(5):823-40. Epub 2016/05/19. doi: 10.1111/mmi.13425. PubMed
533 PMID: 27191515.
534 41. Papenfort K, Forstner KU, Cong JP, Sharma CM, Bassler BL. Differential RNA-seq of Vibrio
535 cholerae identifies the VqmR small RNA as a regulator of biofilm formation. Proceedings of
536 the National Academy of Sciences of the United States of America. 2015;112(7):E766-75. Epub
537 2015/02/04. doi: 10.1073/pnas.1500203112. PubMed PMID: 25646441; PMCID:
538 PMC4343088.
539 42. Shao Y, Bassler BL. Quorum regulatory small RNAs repress type VI secretion in Vibrio
540 cholerae. Mol Microbiol. 2014;92(5):921-30. Epub 2014/04/05. doi: 10.1111/mmi.12599.
541 PubMed PMID: 24698180; PMCID: PMC4038675.
542 43. Drebes Dörr NC, Proutière A, Jaskólska M, Stutzmann S, Bader L, Blokesch M. A single

543 nucleotide polymorphism determines constitutive versus inducible type VI secretion in
544 Vibrio cholerae. bioRxiv. 2022:2022.01.28.478222. doi:
545 10.1101/2022.01.28.478222.
546 44. Gourse RL, Ross W, Gaal T. UPs and downs in bacterial transcription initiation: the role of
547 the alpha subunit of RNA polymerase in promoter recognition. Mol Microbiol.
548 2000;37(4):687-95. Epub 2000/09/06. doi: 10.1046/j.1365-2958.2000.01972.x. PubMed
549 PMID: 10972792.
550 45. Guo C, McDowell IC, Nodzenski M, Scholtens DM, Allen AS, Lowe WL, Reddy TE.
551 Transversions have larger regulatory effects than transitions. BMC Genomics. 2017;18(1):394.
552 doi: 10.1186/s12864-017-3785-4.
553 46. Slattery M, Zhou T, Yang L, Dantas Machado AC, Gordân R, Rohs R. Absence of a simple
554 code: how transcription factors read the genome. Trends in Biochemical Sciences.
555 2014;39(9):381-99. doi: 10.1016/j.tibs.2014.07.002.
556 47. MacKenzie KD, Wang Y, Musicha P, Hansen EG, Palmer MB, Herman DJ, Feasey NA, White
557 AP. Parallel evolution leading to impaired biofilm formation in invasive Salmonella strains.
558 PLoS Genet. 2019;15(6):e1008233. Epub 2019/06/25. doi: 10.1371/journal.pgen.1008233.
559 PubMed PMID: 31233504; PMCID: PMC6611641.
560 48. Osborne SE, Walthers D, Tomljenovic AM, Mulder DT, Silphaduang U, Duong N, Lowden
561 MJ, Wickham ME, Waller RF, Kenney LJ, Coombes BK. Pathogenic adaptation of intracellular

562 bacteria by rewiring a cis-regulatory input function. Proceedings of the National
563 Academy of Sciences. 2009;106(10):3982-7. doi: 10.1073/pnas.0811669106.
564 49. Unterweger D, Kitaoka M, Miyata ST, Bachmann V, Brooks TM, Moloney J, Sosa O, Silva D,
565 Duran-Gonzalez J, Provenzano D, Pukatzki S. Constitutive type VI secretion system expression
566 gives Vibrio cholerae intra- and interspecific competitive advantages. PLoS One.
567 2012;7(10):e48320. Epub 2012/10/31. doi: 10.1371/journal.pone.0048320. PubMed PMID:
568 23110230; PMCID: PMC3482179.
569 50. Fu Y, Waldor MK, Mekalanos JJ. Tn-Seq analysis of Vibrio cholerae intestinal colonization
570 reveals a role for T6SS-mediated antibacterial activity in the host. Cell Host Microbe.
571 2013;14(6):652-63. Epub 2013/12/18. doi: 10.1016/j.chom.2013.11.001. PubMed PMID:
572 24331463; PMCID: PMC3951154.
573 51. Zhao W, Caro F, Robins W, Mekalanos JJ. Antagonism toward the intestinal microbiota
574 and its effect on Vibrio cholerae virulence. Science. 2018;359(6372):210-3. Epub 2018/01/13.
575 doi: 10.1126/science.aap8775. PubMed PMID: 29326272; PMCID: PMC8010019.
576 52. Logan SL, Thomas J, Yan J, Baker RP, Shields DS, Xavier JB, Hammer BK, Parthasarathy R.
577 The Vibrio cholerae type VI secretion system can modulate host intestinal
578 mechanics to displace gut bacterial symbionts. Proceedings of the National Academy of
579 Sciences. 2018;115(16):E3779-E87. doi: 10.1073/pnas.1720133115.
580 53. Seal S, Dharmarajan G, Khan I. Evolution of pathogen tolerance and emerging infections:

581 A missing experimental paradigm. Elife. 2021;10. Epub 2021/09/22. doi: 10.7554/eLife.68874.

582 PubMed PMID: 34544548; PMCID: PMC8455132.

583 54. San JE, Baichoo S, Kanzi A, Moosa Y, Lessells R, Fonseca V, Mogaka J, Power R, de Oliveira

584 T. Current Affairs of Microbial Genome-Wide Association Studies: Approaches, Bottlenecks

585 and Analytical Pitfalls. Front Microbiol. 2019;10:3119. Epub 2020/02/23. doi:

586 10.3389/fmicb.2019.03119. PubMed PMID: 32082269; PMCID: PMC7002396.

587 55. Skorupski K, Taylor RK. Positive selection vectors for allelic exchange. Gene.

588 1996;169(1):47-52. Epub 1996/02/22. doi: 10.1016/0378-1119(95)00793-8. PubMed PMID:

589 8635748.

590 56. Watve SS, Bernardy EE, Hammer BK. Vibrio cholerae: Measuring Natural Transformation

591 Frequency. Curr Protoc Microbiol. 2014;35:6A 4 1-12. Epub 2014/11/05. doi:

592 10.1002/9780471729259.mc06a04s35. PubMed PMID: 25367272.

593 57. Matson JS. Preparation of Vibrio cholerae Samples for RNA-seq Analysis. Methods Mol

594 Biol. 2018;1839:29-38. Epub 2018/07/27. doi: 10.1007/978-1-4939-8685-9_3. PubMed PMID:

595 30047051.

596 58. Manera K, Caro F, Li H, Pei TT, Hersch SJ, Mekalanos JJ, Dong TG. Sensing of intracellular

597 Hcp levels controls T6SS expression in Vibrio cholerae. Proceedings of the National Academy

598 of Sciences of the United States of America. 2021;118(25). Epub 2021/06/24. doi:

599 10.1073/pnas.2104813118. PubMed PMID: 34161288; PMCID: PMC8237632.

600 59. Sayers EW, Agarwala R, Bolton EE, Brister JR, Canese K, Clark K, Connor R, Fiorini N, Funk
601 K, Hefferon T, Holmes JB, Kim S, Kimchi A, Kitts PA, Lathrop S, Lu Z, Madden TL, Marchler-Bauer
602 A, Phan L, Schneider VA, Schoch CL, Pruitt KD, Ostell J. Database resources of the National
603 Center for Biotechnology Information. Nucleic Acids Res. 2019;47(D1):D23-D8. Epub
604 2018/11/06. doi: 10.1093/nar/gky1069. PubMed PMID: 30395293; PMCID: PMC6323993.
605 60. Camacho C, Coulouris G, Avagyan V, Ma N, Papadopoulos J, Bealer K, Madden TL. BLAST+:
606 architecture and applications. BMC Bioinformatics. 2009;10:421. Epub 2009/12/17. doi:
607 10.1186/1471-2105-10-421. PubMed PMID: 20003500; PMCID: PMC2803857.
608 61. Edgar RC. MUSCLE: multiple sequence alignment with high accuracy and high throughput.
609 Nucleic Acids Res. 2004;32(5):1792-7. Epub 2004/03/23. doi: 10.1093/nar/gkh340. PubMed
610 PMID: 15034147; PMCID: PMC390337.
611 62. Madeira F, Park YM, Lee J, Buso N, Gur T, Madhusoodanan N, Basutkar P, Tivey ARN,
612 Potter SC, Finn RD, Lopez R. The EMBL-EBI search and sequence analysis tools APIs in 2019.
613 Nucleic Acids Res. 2019;47(W1):W636-W41. Epub 2019/04/13. doi: 10.1093/nar/gkz268.
614 PubMed PMID: 30976793; PMCID: PMC6602479.
615 63. Felsenstein J. Evolutionary trees from DNA sequences: a maximum likelihood approach.
616 J Mol Evol. 1981;17(6):368-76. Epub 1981/01/01. doi: 10.1007/BF01734359. PubMed PMID:
617 7288891.
618 64. Tamura K, Stecher G, Kumar S. MEGA11: Molecular Evolutionary Genetics Analysis

619 Version 11. Mol Biol Evol. 2021;38(7):3022-7. Epub 2021/04/24. doi:

620 10.1093/molbev/msab120. PubMed PMID: 33892491; PMCID: PMC8233496.

621