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[5], 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

[5]Biology Department, Spelman College, Atlanta, Georgia, USA

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

Abstract:

Mutations in regulatory mechanisms that control gene expression contribute to phenotypic diversity and thus facilitate the adaptation of microbes to new niches. Regulatory architecture is often inferred from transcription factor identification and genome analysis using purely computational approaches. However, there are few examples of phenotypic divergence that arise from the rewiring of bacterial regulatory circuity by mutations in intergenic regions, because locating regulatory elements within regions of DNA that do not code for protein

requires genomic and experimental data. We identify a single *cis*-acting single nucleotide polymorphism (SNP) dramatically alters control of the type VI secretion system (T6), a common weapon for inter-bacterial competition. Tight T6 regulatory control is necessary for adaptation of the waterborne pathogen *Vibrio cholerae* to *in vivo* conditions within the human gut, which we show can be altered by this single non-coding SNP that results in constitutive expression *in vitro*. Our results support a model of pathogen evolution through *cis*-regulatory mutation and preexisting, active transcription factors, thus conferring different fitness advantages to tightly regulated strains inside a human host and unfettered strains adapted to environmental niches.

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 motifs are major contributors to phenotypic diversity (1-3). In bacteria, where pervasive HGT of TFs can alter entire regulatory circuits, CRE mutations receive little attention as drivers of phenotypic divergence (4). 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

39 experimental approaches to map the *cis* and *trans* regulatory interactions across and within

40 species.

41

42 To understand how mutations play a role in microbial adaptation, pathogenic viruses and

43 bacteria with lifestyles that exploit niches within and outside a human host are of great

44 interest. Following ingestion, pandemic strains of the bacterium *Vibrio cholerae* can colonize

45 the human gastrointestinal tract and secrete the cholera toxin that leads to the often fatal

46 diarrhea responsible for seven pandemics to date (5-7). Conversely, *V. cholerae* isolated from

47 non-human niches lack the horizontally-acquired prophage that carries the cholera toxin, and

48 cause mild illness (8). By contrast, all sequenced *V. cholerae* encode a Type VI Secretion System

49 (T6), a broadly distributed "nano-harpoon" weapon that injects into neighboring bacterial

50 cells toxic effector proteins that damage the cell envelope and cause lysis (9, 10). Due to its

51 broad distribution among bacteria including those of the human gut, there is intense interest

52 in understanding the T6 interactions between our microbiota and foreign pathogens, and

53 whether they can be manipulated to influence health (11).

54

55 *V. cholerae* obtained from humans carry a limited T6 arsenal of effectors believed to be

56 tailored for *in vivo* success (9, 12-16), strains from non-human niches encode a more diverse

57 effector repertoire (9, 12, 17, 18). To date, however, adaptative evolution mechanisms of T6

58 regulation in *V. cholerae* derived from non-human sources have largely been overlooked.
59 Studies of human-derived strains identify two primary TFs for T6 activation (19-23). T6 control
60 in pandemic strains (e.g. C6706) requires QstR (Quorum-Sensing and TfoX-Dependent
61 Regulator), which integrates multiple external cues (24-27), and contains a DNA binding
62 domain postulated to interact with a presumptive CRE of the major T6 gene cluster (20, 24).
63 T6 regulation in non-pandemic strain V52, which causes mild disease, requires TfoY,
64 modulatable by intracellular signals (22, 23). How QstR and TfoY control T6 transcription
65 remains elusive, with no T6 CRE yet described. Elucidation of the differences in intraspecies
66 T6 regulatory mechanisms between diverse *V. cholerae* isolates will provide insights into how
67 pathogens emerge from nonpathogenic progenitor.
68
69 To understand the regulatory differences in *V. cholerae* strains, we examine here several
70 environmental isolates that exhibit T6-mediated killing (28). Despite encoding functional
71 signaling circuity and TFs, we find that QstR is dispensable for killing with TfoY playing only a
72 minor role in the strains tested. Thus, existing regulatory models fail to explain the T6 control
73 in *V. cholerae* from human and non-human sources. Genomic analysis identifies one
74 conserved non-coding single-nucleotide polymorphism (SNP) that we show interconverts *V.*
75 *cholerae* T6 activity between chitin-inducible and constitutive states, which are QstR-
76 dependent and TfoY-independent, respectively. We demonstrate that non-coding SNPs can

77 rewire cis-regulatory elements to aid in adaptation of bacteria to different niches, including

78 the human host.

79

80 **Results and Discussion**

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

82 In the absence of chitin, pandemic C6706 with *qstR* expressed from a heterologous promoter

83 (*qstR**) reduces survival of *Escherichia coli* "target" cells in co-culture by over 4-orders of

84 magnitude (~10,000), compared to wildtype (WT) C6706, T6^{-} (*ΔvasK)*, and *ΔqstR* mutants (Fig.

85 1A) (26). Deletion of *tfoY* does not reduce the killing activity of the T6^{+} *qstR** strain, but

86 eliminates the robust killing in non-pandemic strain V52 (serogroup O37), which does not

87 require QstR (Fig. 1B) (23).

88

89 To determine whether either QstR or TfoY participates in control of the T6 in non-human

90 strains, we examined 3223-74, a genetically-amenable, T6-proficient environmental strain

91 (28). Like V52, 3223-74 does not require QstR to efficiently kill *E. coli* in conditions without

92 chitin, but surprisingly, also does not require TfoY. The *ΔtfoY* and *ΔqstR ΔtfoY* mutants retain

93 >99.99% killing activity, with only modest *E. coli* survival (Fig. 1C). Gene fusions of the 5'

94 intergenic region (IGR) of the major T6 cluster of each strain fused to green fluorescent protein

95 (*gfp*) confirm that transcriptional differences account for the killing observed, with maximal

96 *gfp* expression mirroring activity (i.e. low *E. coli* survival with high *gfp* expression, and *vice*

97 *versa*) (Fig. 1D-F). Confocal microscopy reinforces the negligible role of TfoY on killing by 3223-

98 74 (Fig. 1G). Transcription of plasmid-borne reporters is significantly higher in *V. cholerae* than

99 in *E. coli* (Fig. S1), supporting a hypothesis that an additional *V. cholerae*-specific regulator of

100 the T6 likely remains to be identified (Fig. S1).

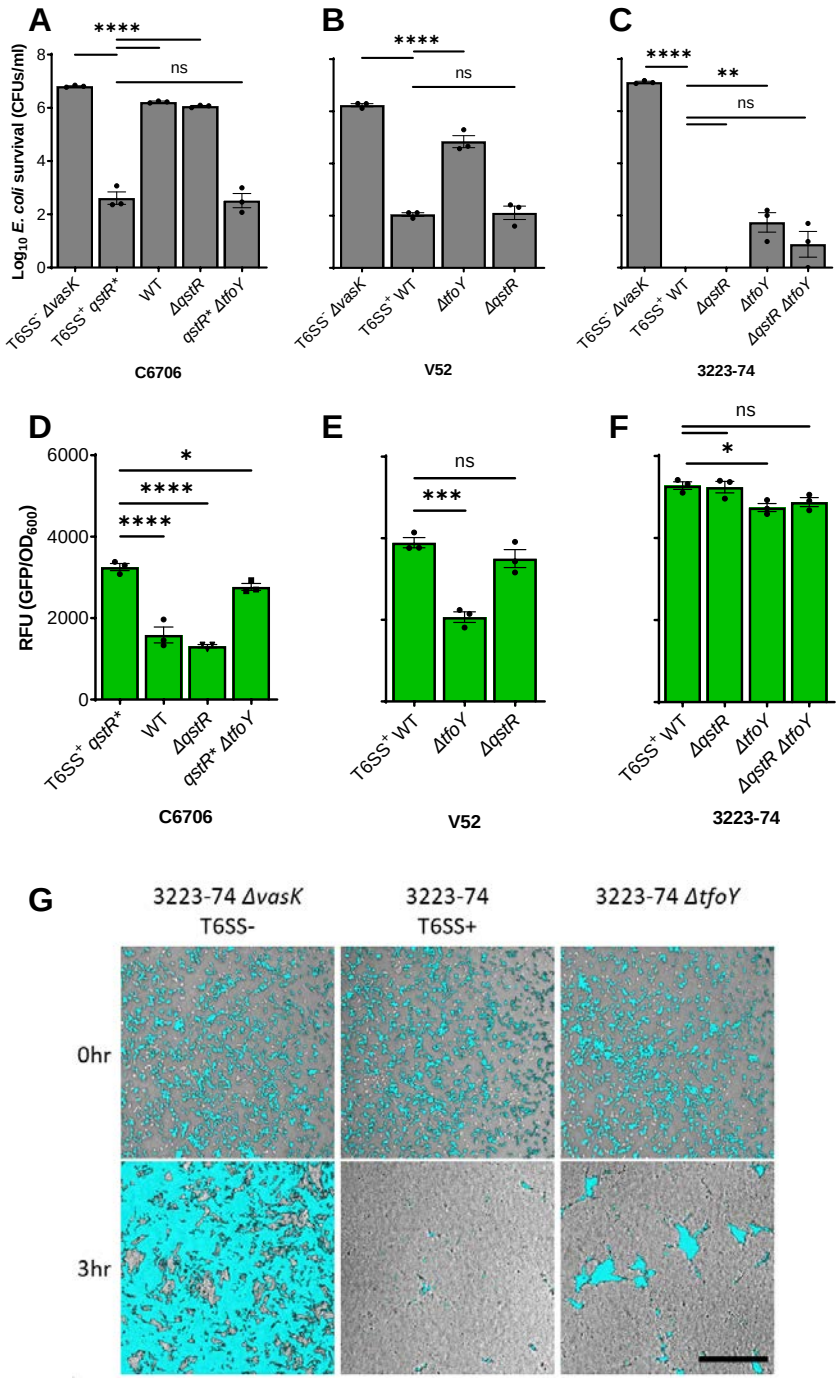


Figure 1. *Vibrio cholerae* 3223-74 T6 activity is QstR- and TfoY-independent. (A-C) Indicated *V. cholerae* strains were co-cultured with chloramphenicol resistant (Cm^R) *E. coli* followed by determination of *E. coli* survival by counting of colony forming units (CFUs) on LB agar with Cm. (D-F) Fluorescence levels are from reporters with *gfp* fused to the intergenic region 5' of

106 *vipA* derived from the three strains shown. The mean value ± S.E. of three independent co-
107 cultures (A-C) and monocultures (D-F) are shown from one experiment, with similar results
108 obtained in at least two other independent experiments. A one-way ANOVA with Dunnett
109 post-hoc test was conducted to determine the significance: ns denotes not significant, ****p
110 ≤ 0.0001, ***p ≤ 0.001, **p ≤ 0.01, *p ≤ 0.05. (G) *E. coli* cells expressing constitutive *gfp* were
111 competed against 3223-74, with the same frame imaged at 0 h and 3 h by confocal microscopy.
112 In the images, *gfp* signal from the *E. coli* is overlaid on top of bright-light images of the co-
113 culture. Scale bar = 50 µm.
114
115 To probe each strain's T6-related regulatory circuitry, we measured canonical behaviors under
116 control of HapR, QstR and TfoY; quorum sensing (QS) controlled bioluminescence, natural
117 transformation, and motility, respectively (29-31). As expected, each TF is intact in C6706, and
118 V52 is QS^- due to a *hapR* mutation (32) but still encodes a functional *tfoY* that controls motility
119 (Fig. S2A-B) (33). Interestingly, the regulatory circuity of *V. cholerae* 3223-74 is intact, like
120 C6706, confirming that it encodes functional TFs (Fig. S2C) that are expendable for T6-
121 mediated killing.
122
123 A SNP in the T6 intergenic region confers QstR-dependency.
124 Human and environmental isolates of *V. cholerae* we have characterized prior (28) share ≥97%

125 average nucleotide identity with many chromosomal differences (9), but inspection of the T6

126 IGRs of C6706, V52 and 3223-74 revealed only 17 SNPs and 3 multinucleotide polymorphisms

127 (Fig. 2A), which we hypothesized could contribute to the differences in transcription and killing

128 activity observed. To address this, we replaced the T6 IGR of C6706 on the chromosome with

129 that from V52 and 3223-74 and measured killing activity. While C6706 carrying the *qstR** allele,

130 but not WT, can adeptly kill *E. coli*, both IGR replacements increase the killing efficiency of WT

131 C6706 by 5- to 6-orders of magnitude (Fig. 2B), mimicking the robust killing observed by WT

132 V52 and 3223-74 (Fig. 1B-C). Deletion of *tfoY* in C6706 with V52's IGR increases *E. coli* survival

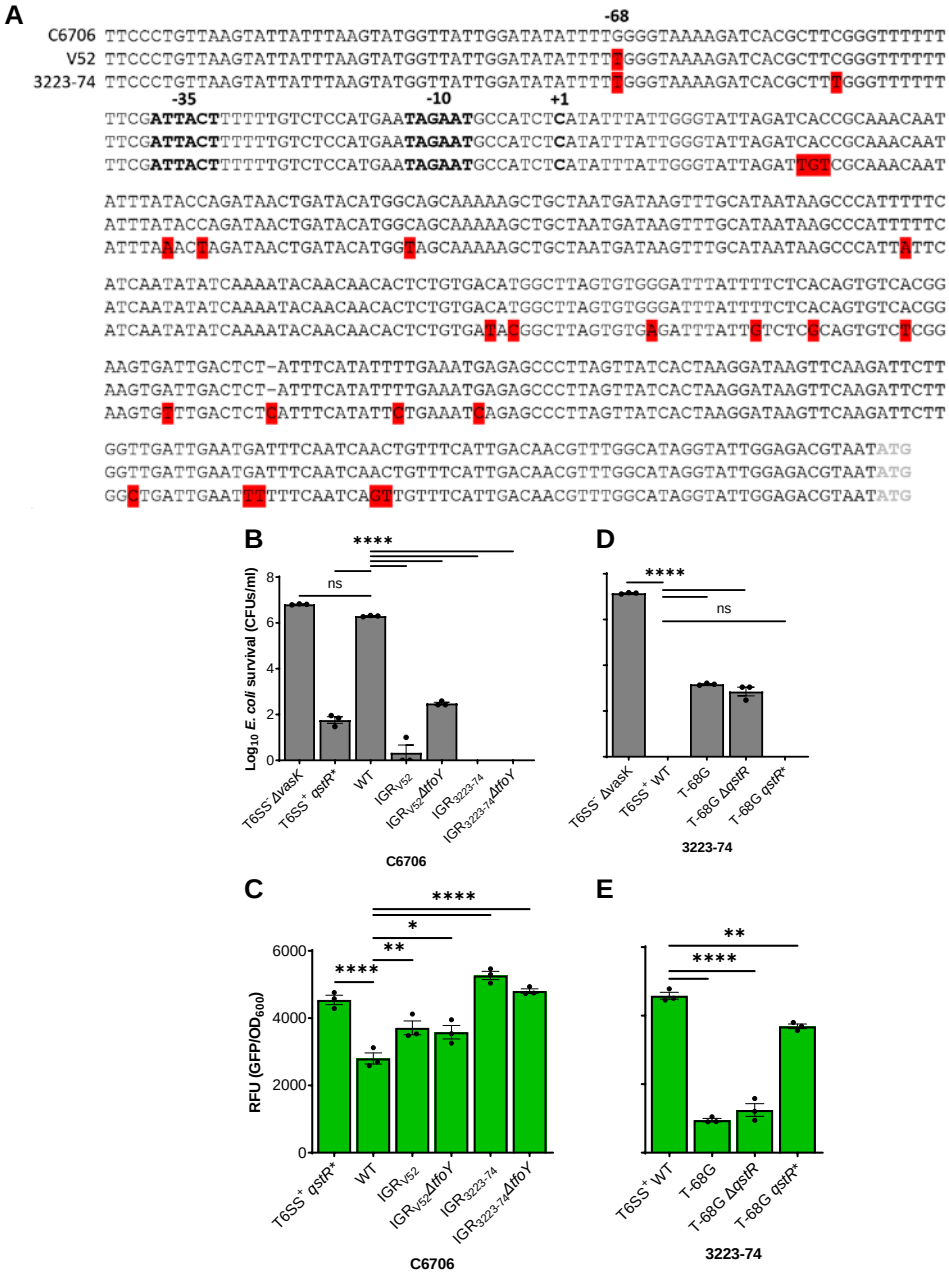
133 (~ 2-logs), as observed with V52, but does not alter *E. coli* survival with 3223-74's IGR (Fig. 2B).

134 Chromosomal transcriptional *gfp* reporters with identical mutations were elevated relative to

135 WT C6706 in each IGR replacement strain (Fig. 2C), consistent with the enhanced killing

136 detected. These results support a hypothesis that a novel CRE lies within the IGR 5' of the T6

137 locus, despite a lack of any known direct TFs of this locus identified to date.



138

139 **Figure 2. G-388T mutation abolishes QstR dependence in C6706 and T-388G confers QstR**

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

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

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

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

144 A T-68G mutation in the 5' IGR of *vipA* was introduced into 3223-74 with different *qstR* alleles.

145 Competition assays were conducted by co-culturing *V. cholerae* killers and Cm^R *E. coli* target

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

147 agar with Cm. (C, E) Shown are fluorescence levels of transcriptional reporters with *gfp* fused

148 to corresponding IGRs of *vipA* expressed in either C6706 (C) or 3223-74 (E). The mean value ±

149 S.E. of three independent co-cultures (B and D) and monocultures (C and E) are shown from

150 one experiment, with similar results were obtained in at least two other independent

151 experiments. A one-way ANOVA with Dunnett post-hoc test was conducted to determine the

152 significance - ns: not significant, ****$p \leq 0.0001$, **$p \leq 0.01$, *$p \leq 0.05$.

153

154 To begin mapping the T6 IGR region and SNP locations, we experimentally determined

155 (Methods) the transcriptional start site (+1) 320 nucleotides (nt) 5' of the ATG of the first T6

156 gene (*vipA*), and adjacent to a putative promoter with 8/12 identical nts compared to the

157 consensus for housekeeping genes (Fig. 2A). The +1 is consistent with paired-end RNAseq

158 results (26), and 320 nt 5' untranslated region (UTR) suggesting possible post-transcriptional

159 regulation, beyond the sRNA interactions already described near the ribosome binding site

160 (RBS) (34). Alignment of the IGRs of C6706 and V52 reveals a single SNP at -68, with a guanine

161 (G) in C6706 at that position and a thymine (T) in V52 (Fig. 2A). Because the replacement of

162 the C6706 IGR with V52 was effectively a G-68T mutation (Fig. 2B-C), we further tested

163 whether G was necessary for QstR activation by replacing the T with a G at position -68 (T-
164 68G) in the 3223-74 WT, *qstR**, and *ΔqstR* backgrounds. The T-68G mutation significantly
165 increases *E. coli* survival and T6 transcription in WT 3223-74 and the *ΔqstR* mutant, with killing
166 restored in the *qstR** background (Fig. 2D-E). Thus, a G at position -68 confers QstR-control,
167 while a T results in constitutive killing *in vitro*.
168
169 The SNP at -68 is evolutionarily conserved.
170 To determine the prevalence of the SNP at -68 in *V. cholerae* strains, we aligned the T6 IGR
171 sequences of diverse strains that exhibit T6 killing (Fig. 3A) (28). Consistent with prior studies
172 (9, 12, 14, 16), our phylogenetic analysis of the T6 IGRs places human strains in a distinct clade,
173 with the exception of two O1 strains isolated nearly a century ago (NCTC8457 and MAK757),
174 and two non-O1 strains (MZO-2 O14 and V52 O37; Fig. S3). All 23 environmental isolates carry
175 the T-68 SNP and display constitutive T6, with one exception that is chitin-inducible (1496-86)
176 (Fig. 3A, S5). By contrast, the 18 human isolates tested carry either G or T at the -68 position
177 (Fig. 3A, S6). The 13 chitin-inducible human isolates carry a G; five show constitutive activity
178 and carry a T like environmental strains, with one exception that is constitutive yet carries the
179 G (2010EL-1749) (Fig. 3A, S6). Neither C nor A are observed at -68 in any stains tested,
180 although both pyrimidine nucleotides (T and C) confer constitutive killing at -68, and both
181 purines (G and A) behave similarly (Fig. S4). The focal SNP location is distal from the -35,

182 inconsistent with an AT-rich "UP element" (35) and more likely a component of a CRE for a TF

183 to be determined. Indeed, transversion mutations have greater effects of TF binding than

184 transitions, as noted here (Fig. S4) likely due to changes in shape of the DNA backbone or

185 DNA-amino acid contacts (36, 37).

186

187 We examined regulation of three additional manipulatable environmental strains (VC22,

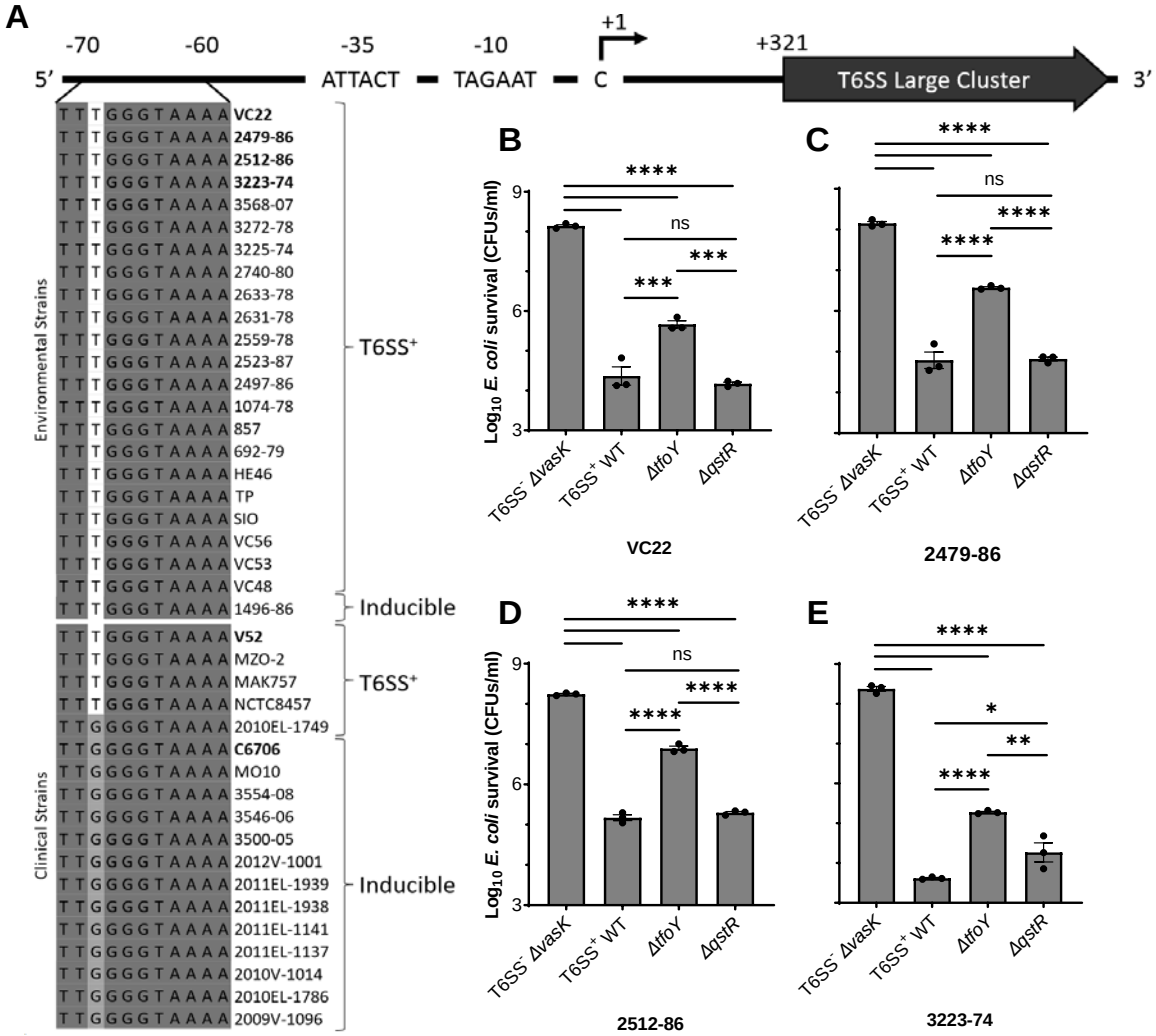
188 2479-89, and 2512-86) that exhibit T6 killing (28). Like 3223-74, qstR is expendable in each

189 strain (Fig. 3B-E). TfoY contributes to some extent in activating T6, with varying *E. coli* recovery

190 observed in each *ΔtfoY* mutant (Fig. 3B-E). Taken together, our findings reveal that the

191 constitutive T6 killing activity of environmental *V. cholerae* is driven by a T at position -68,

192 which obviates the QstR requirement, and permits modest TfoY regulation.



193

194 **Figure 3. Environmental *V. cholerae* isolates encode a T at position -68 while clinical chitin-**

195 **induced isolates encode a G.** (A) A SNP at position -68 in the IGR of the major T6 cluster

196 controls killing activity. Conserved nts are in dark grey and the SNP of interest is highlighted in

197 white/grey. T6 control was categorized as described (28). (B-E) Survival of *E. coli* following

198 competition assays with WT *V. cholerae* strains and mutants was determined by CFU counts.

199 Data shown are the mean ± S.E. of three independent experiments. A one-way ANOVA with

200 Tukey post-hoc test was conducted to determine the significance - ns: not significant, ****p ≤

201 0.0001, ***p ≤ 0.001, **p ≤ 0.01, *p ≤ 0.05.

202

203 While T6s were first described in *V. cholerae* in 2006, the identity of a TF that directly controls

204 the major T6 cluster has remained elusive (19, 21). We speculate that the focal SNP at position

205 -68 is a component of a CRE that contributes to pathoadaptation (Fig. 2A), with T6 activity

206 constrained by QstR in hosts devoid of chitin signals (38). Indeed, dramatic changes in host

207 mechanics induced by constitutive QstR expression may be detrimental to sustained host

208 association (39). By contrast, the T6 IGRs of environmental strains carry a T at -68 and

209 additional SNPs (Fig. S5, S6) that may allow TfoY contributions at the promoter DNA or in the

210 5' UTR (Fig. 2A and Fig. 3B-E). Microbial genome wide associations studies hold great promise

211 for identifying focal SNPs (40), with our study showcasing the power of combining genomic,

212 genetic and phenotypic analysis to identify polymorphisms contributing to pathogen

213 evolution.

214

215 **Materials and Methods**

216 Bacterial growth conditions and plasmid constructions

217 All *V. cholerae* and *E. coli* (Table S1) strains were grown aerobically at 37 °C overnight in

218 Lysogeny Broth (LB) with constant shaking or statically on LB agar. Ampicillin (100 µg/ml),

219 kanamycin (50 µg/ml), chloramphenicol (10 µg/ml), spectinomycin (100 µg/ml), streptomycin

220 (5 mg/ml), sucrose (20% w/v) and diaminopimelic acid (50 µg/ml) were supplemented where

221 appropriate.

222

223 Plasmids (Table S2) used were constructed with DNA restriction nucleases (Promega – WI,

224 USA), Gibson Assembly mix (New England Biolabs – MA, USA), and PCR amplification (Qiagen

225 - Hilden, Germany) by PCR with Q5 polymerase (New England Biolabs – MA, USA), and primers

226 (Table S3) generated by Eton Bioscience Inc (NC, USA). All reagents were used according to

227 the manufacturer's instructions. Plasmids were confirmed by PCR and Sanger sequencing by

228 Eton Bioscience Inc (NC, USA).

229

230 *V. cholerae* mutant construction

231 All genetically engineered strains of *V. cholerae* were constructed with established allelic

232 exchange methods using vector pKAS32 (41) and pRE118 (Addgene - Plasmid #43830). All

233 Insertions, deletions, and mutations were confirmed by PCR and Sanger sequencing

234 conducted by Eton Bioscience Inc (NC, USA). Primers used are in Table S3.

235

236 Fluorescence microscopy

237 *V. cholerae* 3223-74 strains and chromosomal-labeled GFP *E. coli* were separately back-diluted

238 1:100 and incubated at 37 °C for 3 h. *V. cholerae* and *E. coli* were normalized to OD_{600} = 1 and

239 mixed in a 1:5 ratio. A 2 µL aliquot of a mixed culture was inoculated on LB agar and allowed

240 to dry. Cells were imaged before and after a 3 h incubation at 37 °C and 96-100% humidity

241 using an Eclipse Ti-E Nikon (NY, USA) inverted microscope with a Perfect Focus System and

242 camera previously described (9). The images were processed with ImageJ (31).

243

244 Motility assay

245 Overnight cultures of *V. cholerae* were diluted to OD_{600} = 0.1, and 1 µL inoculated onto pre-

246 dried LB plates with 0.3 % agar. Cells were incubated at 37 °C statically overnight, with motility

247 determined by measuring the swarming diameter.

248

249 Transformation assay

250 Chitin-induced transformation frequency was measured as described with defined artificial

251 sea water (450 mM NaCl, 10 mM KCl, 9 mM $CaCl_2$, 30 mM $MgCl_2 \cdot 6H_2O$, and 16 mM

252 $MgSO_4 \cdot 7H_2O$; pH 7.8) (42). Bacteria were incubated with extracellular DNA in triplicate wells

253 containing crab shell tabs, and transformation frequency calculated as Spectinomycin

254 resistant ($Spec^R$) CFU ml^{-1} / total CFU ml^{-1}.

255

256 QS-dependent Luciferase assay

257 Overnight cultures of the bacterial strains were diluted to OD_{600} = 0.001 in liquid LB in

258 microtiter plates and incubated at 37 °C with shaking. The OD_{600} and luminescence were

259 measured each h with a BioTek (VT, USA) Synergy H1 microplate reader to calculate Relative

260 Luminescence Units (RLU) as Luminescence/OD_{600}. Data were collected when OD_{600} = 0.6-0.8.

261 LB medium was used as the blank for the OD_{600} and luminescence.

262

263 GFP transcriptional reporter quantification

264 Overnight cultures of *V. cholerae* or *E. coli* were diluted 1:100 and incubated at 37 °C for 3 h.

265 To enhance the translation of *gfp*, the sequence of the native RBS (12 nt sequence) was

266 replaced with the T7 RBS (12 nt sequence) in the primers used to make the fusions. Cm was

267 added to maintain the plasmid-borne versions of reporters that were cloned into plasmid

268 pSLS3. 300 µL aliquots were transferred to black microtiter plates to read the OD_{600} and GFP

269 fluorescence (Excitation: 485, Emission: 528) with a BioTek Synergy H1 microplate reader (VT,

270 USA) to calculate Relative Luminescence Units (RLU) as Luminescence/OD_{600}. LB medium was

271 used as the blank for the OD_{600}. Strain lacking reporters served as blanks for GFP fluorescence.

272 RFU was calculated by blanked GFP fluorescence / blanked OD_{600}.

273

274 T6-mediated killing assay

275 Overnight cultures of *V. cholerae* or *E. coli* were back-diluted 1:100 and incubated at 37 °C for

276 3 h. *V. cholerae* strains and the Cm^R *E. coli* target were normalized to OD_{600} = 1 and then mixed

277 at a ratio of either 10:1 or 1:5. A 50 µL mixed culture was spotted onto LB agar and dried. After
278 a 3 h incubation at 37°C, cells were resuspended in 5 ml of LB, and serial dilutions were
279 conducted. Finally, the resuspension was inoculated on a LB agar containing Cm to select for
280 the surviving *E. coli*, which was incubated overnight at 37 °C and the *E. coli* colonies were
281 counted and shown as CFU mL^{-1}.
282
283 RNA extraction and determination of the +1 of transcription by 5'-RACE
284 Overnight cultures of *V. cholerae* were back-diluted 1:100 and incubated at 37 °C for 3 h before
285 lysing. Three independent cultures of T6-active *V. cholerae* C6706 *qstR** and 3223-74 WT were
286 harvested by centrifugation at room temperature. RNA isolation, genomic DNA removal, and
287 RNA clean-up were performed as previously described (43). Genomic DNA contamination was
288 confirmed by conducting PCR with primer pair specific for 16S rRNA loci (*rrsA*) as previously
289 described (Table S3) (44). RNA purity was confirmed by NanoDrop (260 / 280 ≈ 2.0).
290
291 5'-RACE (Invitrogen™ - MA, USA) was conducted according to the manufacturer's protocol
292 with slight modifications. Specifically, SuperScript™ IV reverse transcriptase (Invitrogen™ -
293 MA, USA) was used to complete the first strand cDNA synthesis. Two *vipA*-specific primers
294 (GT3056 and GT3060) were used to identify the +1 of transcription for the major T6 gene
295 cluster (Table S3). PCR products were purified with QIAquick PCR purification kit (Qiagen -

296 Hilden, Germany) or Zymoclean gel DNA recovery kit (Zymo Research - CA, USA). Sanger
297 sequencing was conducted by Eton Bioscience Inc. (NC, USA) with the corresponding nesting
298 primer (Table S3).
299
300 Genomic and phylogenetic analysis
301 Genome sequences of *V. cholerae* strains were collected from NCBI Genome database (Table
302 S4) (45). The IGR upstream of major T6 cluster was extracted, aligned, and presented using
303 BLAST+ v2.2.18 (46), MUSCLE v3.8 (https://www.ebi.ac.uk/Tools/msa/muscle/) (47, 48), and
304 ESPript 3.0 (https://espript.ibcp.fr/) (33). The DNA sequence of the IGR was used for
305 phylogenetic analysis, and the phylogenetic tree was constructed by the Maximum likelihood
306 method using MEGA11 (49, 50).
307
308 For 2012V-1001, 2011EL-1939, 2011EL-1938, and 2011EL-1141 that do not have genome
309 sequence available, colony PCR was conducted to amplify the 5' IGR of the major T6 cluster
310 using OneTaq DNA Polymerase (New England Biolabs – MA, USA). PCR products were
311 confirmed with gel electrophoresis and Sanger sequencing by Eton Bioscience Inc. (NC, USA)
312 with the identical primer pair (Table S3).
313
314 **Acknowledgements**

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

318

319 **Competing interests**

320 The authors have no competing interests.

321 **References**

322 1. Wray GA. The evolutionary significance of cis-regulatory mutations. Nature reviews
323 Genetics. 2007;8(3):206-16. Epub 2007/02/17. doi: 10.1038/nrg2063. PubMed PMID:
324 17304246.

325 2. Stern DL, Orgogozo V. THE LOCI OF EVOLUTION: HOW PREDICTABLE IS GENETIC
326 EVOLUTION? Evolution. 2008;62(9):2155-77. doi: https://doi.org/10.1111/j.1558-
327 5646.2008.00450.x.

328 3. Wittkopp PJ, Kalay G. Cis-regulatory elements: molecular mechanisms and evolutionary
329 processes underlying divergence. Nature reviews Genetics. 2011;13(1):59-69. Epub
330 2011/12/07. doi: 10.1038/nrg3095. PubMed PMID: 22143240.

331 4. Perez JC, Groisman EA. Evolution of transcriptional regulatory circuits in bacteria. Cell.
332 2009;138(2):233-44. Epub 2009/07/28. doi: 10.1016/j.cell.2009.07.002. PubMed PMID:
333 19632175; PMCID: PMC2726713.

334 5. Dziejman M, Balon E, Boyd D, Fraser CM, Heidelberg JF, Mekalanos JJ. Comparative
335 genomic analysis of Vibrio cholerae: genes that correlate with cholera endemic and pandemic
336 disease. Proceedings of the National Academy of Sciences of the United States of America.
337 2002;99(3):1556-61. Epub 2002/01/31. doi: 10.1073/pnas.042667999. PubMed PMID:
338 11818571; PMCID: PMC122229.

339 6. Faruque SM, Albert MJ, Mekalanos JJ. Epidemiology, Genetics, and Ecology of

340 ToxigenicVibrio cholerae. Microbiology and Molecular Biology Reviews.
341 1998;62(4):1301-14. doi: doi:10.1128/MMBR.62.4.1301-1314.1998.
342 7. Karaolis DKR, Johnson JA, Bailey CC, Boedeker EC, Kaper JB, Reeves PR. A Vibrio
343 cholerae pathogenicity island associated with epidemic and pandemic strains.
344 Proceedings of the National Academy of Sciences. 1998;95(6):3134-9. doi:
345 10.1073/pnas.95.6.3134.
346 8. Pruzzo C, Vezzulli L, Colwell RR. Global impact of Vibrio cholerae interactions with chitin.
347 Environ Microbiol. 2008;10(6):1400-10. Epub 2008/03/04. doi: 10.1111/j.1462-
348 2920.2007.01559.x. PubMed PMID: 18312392.
349 9. Crisan CV, Chande AT, Williams K, Raghuram V, Rishishwar L, Steinbach G, Watve SS,
350 Yunker P, Jordan IK, Hammer BK. Analysis of Vibrio cholerae genomes identifies new type VI
351 secretion system gene clusters. Genome Biol. 2019;20(1):163. Epub 2019/08/14. doi:
352 10.1186/s13059-019-1765-5. PubMed PMID: 31405375; PMCID: PMC6691524.
353 10. MacIntyre DL, Miyata ST, Kitaoka M, Pukatzki S. The Vibrio cholerae type VI secretion
354 system displays antimicrobial properties. Proceedings of the National Academy of Sciences of
355 the United States of America. 2010;107(45):19520-4. Epub 2010/10/27. doi:
356 10.1073/pnas.1012931107. PubMed PMID: 20974937; PMCID: PMC2984155.
357 11. Gallegos-Monterrosa R, Coulthurst SJ. The ecological impact of a bacterial weapon:
358 microbial interactions and the Type VI secretion system. FEMS Microbiol Rev. 2021;45(6). Epub

359 2021/06/23. doi: 10.1093/femsre/fuab033. PubMed PMID: 34156081; PMCID: PMC8632748.

360 12. Hussain NAS, Kirchberger PC, Case RJ, Boucher YF. Modular Molecular Weaponry Plays a
361 Key Role in Competition Within an Environmental Vibrio cholerae Population. Frontiers in
362 Microbiology. 2021;12(979). doi: 10.3389/fmicb.2021.671092.

363 13. Ishikawa T, Sabharwal D, Bröms J, Milton DL, Sjöstedt A, Uhlin BE, Wai SN, Camilli A.
364 Pathoadaptive Conditional Regulation of the Type VI Secretion System in Vibrio cholerae O1
365 Strains. Infection and Immunity. 2012;80(2):575-84. doi: doi:10.1128/IAI.05510-11.

366 14. Kirchberger PC, Unterweger D, Provenzano D, Pukatzki S, Boucher Y. Sequential
367 displacement of Type VI Secretion System effector genes leads to evolution of diverse
368 immunity gene arrays in Vibrio cholerae. Scientific Reports. 2017;7(1):45133. doi:
369 10.1038/srep45133.

370 15. Santoriello FJ, Pukatzki S. When the pandemic opts for the lockdown: Secretion system
371 evolution in the cholera bacterium. Microb Cell. 2021;8(3):69-72. Epub 2021/03/09. doi:
372 10.15698/mic2021.03.744. PubMed PMID: 33681340; PMCID: PMC7919388.

373 16. Unterweger D, Miyata ST, Bachmann V, Brooks TM, Mullins T, Kostiuk B, Provenzano D,
374 Pukatzki S. The Vibrio cholerae type VI secretion system employs diverse effector modules for
375 intraspecific competition. Nature Communications. 2014;5(1):3549. doi:
376 10.1038/ncomms4549.

377 17. Crisan CV, Chandrashekar H, Everly C, Steinbach G, Hill SE, Yunker PJ, Lieberman RR,

378 Hammer BK, Ellermeier CD. A New Contact Killing Toxin Permeabilizes Cells and Belongs to a
379 Broadly Distributed Protein Family. mSphere. 2021;6(4):e00318-21. doi:
380 doi:10.1128/mSphere.00318-21.
381 18. Drebes Dörr NC, Blokesch M. Interbacterial competition and anti-predatory behaviour of
382 environmental Vibrio cholerae strains. Environmental Microbiology. 2020;22(10):4485-504.
383 doi: https://doi.org/10.1111/1462-2920.15224.
384 19. Crisan CV, Hammer BK. The Vibrio cholerae type VI secretion system: toxins, regulators
385 and consequences. Environ Microbiol. 2020;22(10):4112-22. Epub 2020/03/07. doi:
386 10.1111/1462-2920.14976. PubMed PMID: 32133757.
387 20. Jaskolska M, Stutzmann S, Stoudmann C, Blokesch M. QstR-dependent regulation of
388 natural competence and type VI secretion in Vibrio cholerae. Nucleic Acids Res.
389 2018;46(20):10619-34. Epub 2018/08/14. doi: 10.1093/nar/gky717. PubMed PMID:
390 30102403; PMCID: PMC6237807.
391 21. Joshi A, Kostiuk B, Rogers A, Teschler J, Pukatzki S, Yildiz FH. Rules of Engagement: The
392 Type VI Secretion System in Vibrio cholerae. Trends Microbiol. 2017;25(4):267-79. Epub
393 2016/12/29. doi: 10.1016/j.tim.2016.12.003. PubMed PMID: 28027803; PMCID:
394 PMC5365375.
395 22. Joshi A, Mahmoud SA, Kim SK, Ogdahl JL, Lee VT, Chien P, Yildiz FH. c-di-GMP inhibits
396 LonA-dependent proteolysis of TfoY in Vibrio cholerae. PLoS Genet. 2020;16(6):e1008897.

397 Epub 2020/06/27. doi: 10.1371/journal.pgen.1008897. PubMed PMID: 32589664; PMCID:
398 PMC7371385.
399 23. Metzger LC, Stutzmann S, Scrignari T, Van der Henst C, Matthey N, Blokesch M.
400 Independent Regulation of Type VI Secretion in Vibrio cholerae by TfoX and TfoY. Cell Rep.
401 2016;15(5):951-8. Epub 2016/04/28. doi: 10.1016/j.celrep.2016.03.092. PubMed PMID:
402 27117415; PMCID: PMC4858559.
403 24. Lo Scrudato M, Blokesch M. A transcriptional regulator linking quorum sensing and chitin
404 induction to render Vibrio cholerae naturally transformable. Nucleic Acids Res.
405 2013;41(6):3644-58. Epub 2013/02/06. doi: 10.1093/nar/gkt041. PubMed PMID: 23382174;
406 PMCID: PMC3616704.
407 25. Meibom KL, Blokesch M, Dolganov NA, Wu CY, Schoolnik GK. Chitin induces natural
408 competence in Vibrio cholerae. Science. 2005;310(5755):1824-7. Epub 2005/12/17. doi:
409 10.1126/science.1120096. PubMed PMID: 16357262.
410 26. Watve SS, Thomas J, Hammer BK. CytR Is a Global Positive Regulator of Competence,
411 Type VI Secretion, and Chitinases in Vibrio cholerae. PLoS One. 2015;10(9):e0138834. Epub
412 2015/09/25. doi: 10.1371/journal.pone.0138834. PubMed PMID: 26401962; PMCID:
413 PMC4581735.
414 27. Tsou AM, Cai T, Liu Z, Zhu J, Kulkarni RV. Regulatory targets of quorum sensing in Vibrio
415 cholerae: evidence for two distinct HapR-binding motifs. Nucleic Acids Res. 2009;37(8):2747-

416 56. Epub 2009/03/12. doi: 10.1093/nar/gkp121. PubMed PMID: 19276207; PMCID:
417 PMC2677876.
418 28. Bernardy EE, Turnsek MA, Wilson SK, Tarr CL, Hammer BK, Liu S-J. Diversity of Clinical and
419 Environmental Isolates of Vibrio cholerae in Natural Transformation and Contact-Dependent
420 Bacterial Killing Indicative of Type VI Secretion System Activity. Applied and Environmental
421 Microbiology. 2016;82(9):2833-42. doi: doi:10.1128/AEM.00351-16.
422 29. Jobling MG, Holmes RK. Characterization of hapR, a positive regulator of the Vibrio
423 cholerae HA/protease gene hap, and its identification as a functional homologue of the Vibrio
424 harveyi luxR gene. Molecular Microbiology. 1997;26(5):1023-34. doi:
425 https://doi.org/10.1046/j.1365-2958.1997.6402011.x.
426 30. Pursley BR, Maiden MM, Hsieh ML, Fernandez NL, Severin GB, Waters CM. Cyclic di-GMP
427 Regulates TfoY in Vibrio cholerae To Control Motility by both Transcriptional and
428 Posttranscriptional Mechanisms. J Bacteriol. 2018;200(7). Epub 2018/01/10. doi:
429 10.1128/JB.00578-17. PubMed PMID: 29311281; PMCID: PMC5847662.
430 31. Schneider CA, Rasband WS, Eliceiri KW. NIH Image to ImageJ: 25 years of image analysis.
431 Nat Methods. 2012;9(7):671-5. Epub 2012/08/30. doi: 10.1038/nmeth.2089. PubMed PMID:
432 22930834; PMCID: PMC5554542.
433 32. Zheng J, Shin OS, Cameron DE, Mekalanos JJ. Quorum sensing and a global regulator TsrA
434 control expression of type VI secretion and virulence in Vibrio cholerae.

435 Proceedings of the National Academy of Sciences. 2010;107(49):21128-33. doi:
436 10.1073/pnas.1014998107.
437 33. Robert X, Gouet P. Deciphering key features in protein structures with the new ENDscript
438 server. Nucleic Acids Res. 2014;42(Web Server issue):W320-4. Epub 2014/04/23. doi:
439 10.1093/nar/gku316. PubMed PMID: 24753421; PMCID: PMC4086106.
440 34. Shao Y, Bassler BL. Quorum regulatory small RNAs repress type VI secretion in Vibrio
441 cholerae. Mol Microbiol. 2014;92(5):921-30. Epub 2014/04/05. doi: 10.1111/mmi.12599.
442 PubMed PMID: 24698180; PMCID: PMC4038675.
443 35. Gourse RL, Ross W, Gaal T. UPs and downs in bacterial transcription initiation: the role of
444 the alpha subunit of RNA polymerase in promoter recognition. Mol Microbiol.
445 2000;37(4):687-95. Epub 2000/09/06. doi: 10.1046/j.1365-2958.2000.01972.x. PubMed
446 PMID: 10972792.
447 36. Guo C, McDowell IC, Nodzenski M, Scholtens DM, Allen AS, Lowe WL, Reddy TE.
448 Transversions have larger regulatory effects than transitions. BMC Genomics. 2017;18(1):394.
449 doi: 10.1186/s12864-017-3785-4.
450 37. Slattery M, Zhou T, Yang L, Dantas Machado AC, Gordân R, Rohs R. Absence of a simple
451 code: how transcription factors read the genome. Trends in Biochemical Sciences.
452 2014;39(9):381-99. doi: 10.1016/j.tibs.2014.07.002.
453 38. Zhao W, Caro F, Robins W, Mekalanos JJ. Antagonism toward the intestinal microbiota

454 and its effect on Vibrio cholerae virulence. Science. 2018;359(6372):210-3. Epub 2018/01/13.
455 doi: 10.1126/science.aap8775. PubMed PMID: 29326272; PMCID: PMC8010019.
456 39. Logan SL, Thomas J, Yan J, Baker RP, Shields DS, Xavier JB, Hammer BK, Parthasarathy R.
457 The Vibrio cholerae type VI secretion system can modulate host intestinal
458 mechanics to displace gut bacterial symbionts. Proceedings of the National Academy of
459 Sciences. 2018;115(16):E3779-E87. doi: 10.1073/pnas.1720133115.
460 40. San JE, Baichoo S, Kanzi A, Moosa Y, Lessells R, Fonseca V, Mogaka J, Power R, de Oliveira
461 T. Current Affairs of Microbial Genome-Wide Association Studies: Approaches, Bottlenecks
462 and Analytical Pitfalls. Front Microbiol. 2019;10:3119. Epub 2020/02/23. doi:
463 10.3389/fmicb.2019.03119. PubMed PMID: 32082269; PMCID: PMC7002396.
464 41. Skorupski K, Taylor RK. Positive selection vectors for allelic exchange. Gene.
465 1996;169(1):47-52. Epub 1996/02/22. doi: 10.1016/0378-1119(95)00793-8. PubMed PMID:
466 8635748.
467 42. Watve SS, Bernardy EE, Hammer BK. Vibrio cholerae: Measuring Natural Transformation
468 Frequency. Curr Protoc Microbiol. 2014;35:6A 4 1-12. Epub 2014/11/05. doi:
469 10.1002/9780471729259.mc06a04s35. PubMed PMID: 25367272.
470 43. Matson JS. Preparation of Vibrio cholerae Samples for RNA-seq Analysis. Methods Mol
471 Biol. 2018;1839:29-38. Epub 2018/07/27. doi: 10.1007/978-1-4939-8685-9_3. PubMed PMID:
472 30047051.

473 44. Manera K, Caro F, Li H, Pei TT, Hersch SJ, Mekalanos JJ, Dong TG. Sensing of intracellular
474 Hcp levels controls T6SS expression in Vibrio cholerae. Proceedings of the National Academy
475 of Sciences of the United States of America. 2021;118(25). Epub 2021/06/24. doi:
476 10.1073/pnas.2104813118. PubMed PMID: 34161288; PMCID: PMC8237632.
477 45. Sayers EW, Agarwala R, Bolton EE, Brister JR, Canese K, Clark K, Connor R, Fiorini N, Funk
478 K, Hefferon T, Holmes JB, Kim S, Kimchi A, Kitts PA, Lathrop S, Lu Z, Madden TL, Marchler-Bauer
479 A, Phan L, Schneider VA, Schoch CL, Pruitt KD, Ostell J. Database resources of the National
480 Center for Biotechnology Information. Nucleic Acids Res. 2019;47(D1):D23-D8. Epub
481 2018/11/06. doi: 10.1093/nar/gky1069. PubMed PMID: 30395293; PMCID: PMC6323993.
482 46. Camacho C, Coulouris G, Avagyan V, Ma N, Papadopoulos J, Bealer K, Madden TL. BLAST+:
483 architecture and applications. BMC Bioinformatics. 2009;10:421. Epub 2009/12/17. doi:
484 10.1186/1471-2105-10-421. PubMed PMID: 20003500; PMCID: PMC2803857.
485 47. Edgar RC. MUSCLE: multiple sequence alignment with high accuracy and high throughput.
486 Nucleic Acids Res. 2004;32(5):1792-7. Epub 2004/03/23. doi: 10.1093/nar/gkh340. PubMed
487 PMID: 15034147; PMCID: PMC390337.
488 48. Madeira F, Park YM, Lee J, Buso N, Gur T, Madhusoodanan N, Basutkar P, Tivey ARN,
489 Potter SC, Finn RD, Lopez R. The EMBL-EBI search and sequence analysis tools APIs in 2019.
490 Nucleic Acids Res. 2019;47(W1):W636-W41. Epub 2019/04/13. doi: 10.1093/nar/gkz268.
491 PubMed PMID: 30976793; PMCID: PMC6602479.

492 49. Felsenstein J. Evolutionary trees from DNA sequences: a maximum likelihood approach.
493 J Mol Evol. 1981;17(6):368-76. Epub 1981/01/01. doi: 10.1007/BF01734359. PubMed PMID:
494 7288891.
495 50. Tamura K, Stecher G, Kumar S. MEGA11: Molecular Evolutionary Genetics Analysis
496 Version 11. Mol Biol Evol. 2021;38(7):3022-7. Epub 2021/04/24. doi:
497 10.1093/molbev/msab120. PubMed PMID: 33892491; PMCID: PMC8233496.