bioRxiv preprint doi: https://doi.org/10.1101/2020.07.24.220178; this version posted July 24, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

Interbacterial competition and anti-predatory behavior of environmental *Vibrio cholerae* strains

Natália C. Drebes Dörr and Melanie Blokesch*

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

*Corresponding author:

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

Running title: Toxicity of environmental *Vibrio cholerae* strains

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

bioRxiv preprint doi: https://doi.org/10.1101/2020.07.24.220178; this version posted July 24, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

Originality-Significance Statement

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

Summary

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

bioRxiv preprint doi: https://doi.org/10.1101/2020.07.24.220178; this version posted July 24, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

44 experimental validations, we showed that highly similar but nonidentical immunity proteins
45 were insufficient to provide cross-immunity among those wild strains.
46
47 **Introduction**
48 Diarrheal diseases can be caused by a variety of microorganisms, including the causative
49 agent of cholera, *Vibrio cholerae*, which infects up to 4 million people every year (Ali *et al.*,
50 2015). Cholera often spreads from its endemic area around the Ganges delta and has reached
51 almost the entire world in the reported seven pandemics that have been witnessed since 1817.
52 In addition to these pandemics, important localized outbreaks have occurred over the years,
53 especially following natural disasters (Faruque *et al.*, 1998; World Health Organization
54 (WHO), 2006; Clemens *et al.*, 2017).
55 *V. cholerae* strains can be classified into serogroups based on more than 200 different
56 O-antigens. The O1 serogroup is considered the primary cause of previous (e.g., classical
57 serotype) and ongoing (e.g., El Tor serotype) cholera pandemics (Faruque *et al.*, 1998;
58 Cottingham *et al.*, 2003). Isolates belonging to the O139 serogroup are genetically related to
59 the 7th pandemic O1 El Tor strains (Johnson *et al.*, 1994) but are rarely associated with
60 disease outbreaks currently (Faruque *et al.*, 1998; Clemens *et al.*, 2017). Even though this
61 species is best known due to the life-threatening disease it causes, the vast majority of *V.*
62 *cholerae* are common members of aquatic habitats. These mostly non-O1/non-O139
63 serogroup strains are thought to frequently associate with zooplankton and shellfish, and *V.*
64 *cholerae* uses their chitinous exoskeletons as a source of carbon and nitrogen (Faruque *et al.*,
65 1998; Cottingham *et al.*, 2003; Kirn *et al.*, 2005). Environmental isolates are typically
66 considered harmless to humans despite reported associations with mild to severe forms of
67 diarrhea (Islam *et al.*, 1992; Singh *et al.*, 2001; Dziejman *et al.*, 2002, 2005; Onifade *et al.*,
68 2011; Deshayes *et al.*, 2015; Hasan *et al.*, 2015).

bioRxiv preprint doi: https://doi.org/10.1101/2020.07.24.220178; this version posted July 24, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

A major question, not only in the cholera field but also for infectious diseases in general, is how pathogenic isolates evolve from their non-pathogenic environmental progenitors. In the case of *V. cholerae*, pandemic patient isolates are remarkably clonal, in sharp contrast to the high genomic variability encountered in environmental isolates (Faruque *et al.*, 1998; Chun *et al.*, 2009; Mutreja *et al.*, 2011; Harris *et al.*, 2012; Domman *et al.*, 2017; Weill *et al.*, 2017, 2019). Two major genetic features common to all pandemic strains are the CTXΦ prophage and the *Vibrio* pathogenicity island (VPI-1 or TCP island). These genetic elements harbor genes encoding the main virulence factors, e.g., cholera toxin (CTX) and toxin-coregulated pilus (TCP) (Taylor *et al.*, 1987; Waldor and Mekalanos, 1996). The possession of these genomic regions, however, is not unique nor a 'deterministic factor' of pandemic strains. In fact, *V. cholerae* samples collected in several regions of the world have exposed environmental CTX- and/or TCP-positive strains (Rivera *et al.*, 2001; Faruque *et al.*, 2003, 2004; Gennari *et al.*, 2012; Bernardy *et al.*, 2016; Shapiro *et al.*, 2016). Additionally, *V. cholerae* strains that have caused localized cholera outbreaks without reaching pandemic levels have been reported, such as the O37 serogroup strains V52 and ATCC25872 (Aldova *et al.*, 1968; Boyd and Waldor, 2002; Cottingham *et al.*, 2003; Chun *et al.*, 2009). Therefore, the terms "toxigenic" (ability to cause cholera) and "pandemic" (here, current 7th pandemic-causing O1 El Tor strains) are not synonymous. What exactly determines whether strains become pandemic or not is still not fully understood and probably involves diverse aspects ranging from genetic content and strain-specific phenotypes, along with the location of the initial outbreaks and ultimately towards social and sanitary factors, to name a few. Pertinent to the bacteria-related aspects, previous studies hypothesized the existence of virulence adaptive polymorphisms (VAP) circulating in environmental strains. These VAPs were suggested to be a prerequisite for pandemic transition before the horizontal acquisition of CTXΦ and VPI-1 (Shapiro *et al.*, 2016).

bioRxiv preprint doi: https://doi.org/10.1101/2020.07.24.220178; this version posted July 24, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

To better understand their emergence, the differences between pandemic strains and their environmental relatives need to be deciphered. Such differences most likely include phenotypic alterations that are not easily predictable by genomics. In the context of phenotypic variation, two minor virulence factors, namely, the type VI secretion system (T6SS) and the pore-forming toxin hemolysin, are of special interest as they are differentially produced in pandemic compared with non-pandemic toxigenic or environmental strains. The T6SS is a molecular killing device that resembles an inverted contractile bacteriophage tail and it is present in approximately 25% of all Gram-negative bacteria (Ho *et al.*, 2014; Cianfanelli *et al.*, 2016; Galán and Waksman, 2018; Taylor *et al.*, 2018). It is composed of a membrane-spanning portion, a tube structure made of stacks of Hcp hexamer rings, and a tube-surrounding sheath. The sheath is composed of the two proteins VipA and VipB which, upon contraction, propels the inner tube out of the cell together with its effector-decorated tip proteins (VgrG and PAAR) (Zoued *et al.*, 2014; Cherrak *et al.*, 2019; Flaugnatti *et al.*, 2020). Secreted Hcp therefore serves as an indicator of T6SS activity (Pukatzki et al 2006; Basler *et al.*, 2012; Bröms *et al.*, 2013; Kube *et al.*, 2014). The secreted effector proteins mostly target conserved cellular components, such as membranes, bacterial peptidoglycan, nucleic acids, or the eukaryotic cytoskeleton (Hood *et al.*, 2010; Russell *et al.*, 2014). T6SS effectors can be either secreted as (i) cargos that interact with T6SS structural proteins, such as Hcp or the tip protein VgrG; or as (ii) C-terminal extensions of VgrG, Hcp, or PAAR (so-called “evolved” proteins; Pukatzki *et al.*, 2006, 2009; Hachani *et al.*, 2016). Notably, T6SS-producing bacteria protect themselves against their toxic effector repertoire by the production of effector-cognate immunity proteins. These effector/immunity (E/I) pairs are usually encoded adjacent to each other (Russell *et al.*, 2011; Brooks *et al.*, 2013; Dong *et al.*, 2013; Unterweger *et al.*, 2014).

bioRxiv preprint doi: https://doi.org/10.1101/2020.07.24.220178; this version posted July 24, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

In current pandemic strains, the T6SS is encoded by four gene clusters: the large cluster and three auxiliary clusters. The large cluster primarily codes for structural proteins (recently reviewed by Crisan and Hammer, 2020), including the evolved tip protein VgrG3, which possesses a C-terminal lysozyme-like domain for peptidoglycan degradation (Zheng *et al.*, 2011; Brooks *et al.*, 2013). Auxiliary clusters 1 and 2 are both composed of genes encoding Hcp and VgrG (evolved in cluster 1 and structural in cluster 2), an adaptor protein (Tap1 or VasW; Liang *et al.*, 2015; Unterweger *et al.*, 2015) and an E/I module. The auxiliary cluster 1 effector TseL is a bifunctional lipase with anti-bacterial and anti-eukaryotic activity (Zheng *et al.*, 2011; Dong *et al.*, 2013; Russell *et al.*, 2013). The auxiliary cluster 2 effector VasX acts as a pore-forming toxin due to its bacterial/eukaryotic colicin-like membrane-disrupting activity (Miyata *et al.*, 2011, 2013; Russell *et al.*, 2014). Finally, the T6SS auxiliary cluster 3 is composed of genes encoding a second copy of PAAR (a tip-sharpening protein that extends from VgrG; the first gene copy is at the start of the large cluster; Shneider *et al.*, 2013) and a single E/I pair. When discovered, the aux 3 effector TseH was predicted to contain a hydrolase domain (Altindis *et al.*, 2015), while the recently reported crystal structure supported its role as a papain-like NlpC/P60 peptidase (Hersch *et al.*, 2020) with structural similarity to the T6SS effector Tse1 of *Pseudomonas aeruginosa* that contains bacteriolytic peptidoglycan amidase activity (Chou *et al.*, 2012).

Another toxic protein that shows differential activity in pandemic compared with non-pandemic *V. cholerae* strains is the hemolysin protein (HlyA), which is likewise widespread among *Vibrio* species (Zhang and Austin, 2005). HlyA is considered a minor virulence factor as it contributes to toxicity in the context of intestinal infections (Ichinose *et al.*, 1987; Olivier *et al.*, 2007). We previously showed that the proper timing of HlyA activity is a prerequisite for pandemic *V. cholerae* to establish a replication niche within the aquatic

bioRxiv preprint doi: https://doi.org/10.1101/2020.07.24.220178; this version posted July 24, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

amoeba *Acanthamoeba castellanii*, while constitutive activity kills this host prematurely (Van der Henst *et al.*, 2018).

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

Results and discussion

Genome sequencing of environmental *V. cholerae* strains

Apart from the major virulence factors, previous observations suggested important phenotypic differences between pandemic and environmental *V. cholerae* strains. We therefore decided to study fifteen environmental isolates (Table S1) from diverse habitats along the central California coast, a region that is free of endemic or epidemic cholera. These strains were initially isolated in 2004 by Keymer and colleagues followed by basic characterization and comparative genome hybridization (CGH) analyses (Keymer *et al.*,

bioRxiv preprint doi: https://doi.org/10.1101/2020.07.24.220178; this version posted July 24, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

2007; Miller *et al.*, 2007). The latter approach was based on amplicon microarrays (representing 3,357 of 3,891 annotated open reading frames), which were designed using the first published *V. cholerae* genome sequence as a template, namely, pandemic O1 El Tor strain N16961 (Heidelberg *et al.*, 2000). Based on the presence or absence of the microarray-templated genes, the strains were classified into four clades (A-D) (Keymer *et al.*, 2007; Miller *et al.*, 2007), a classification we maintained throughout the current report. To better understand the accessory genome, including the T6SS E/I modules and to identify those genes that are novel when compared with strain N16961, we first whole-genome sequenced these strains using a long-read PacBio approach followed by the *de novo* assembly of their genomes. As a representative strain of the 7th cholera pandemic, we used strain O1 El Tor A1552 throughout this study (Yildiz and Schoolnik, 1998). This strain is connected to a cholera outbreak in Peru in the 1990s. We recently reported its genome sequence, including more than 1,000 manual gene annotations, according to previous experimental validations (Matthey *et al.*, 2018).

Table S2 shows the sequencing details and the features of the closed genomes of the 15 environmental isolates. All genomes showed the dual chromosome architecture that is common for *Vibrio* species (Okada *et al.*, 2005), a similar overall size of the two chromosomes, and average GC percentages within the same range as the one observed for the pandemic *V. cholerae* strains N16961 and A1552 (Heidelberg 2000; Matthey *et al.*, 2018). The assembly pipeline also predicted megaplasmids of ~300kbp and 80kbp for four strains belonging to clades C and D, respectively. It should be noted that due to the size selection of the prepared sequencing libraries, putative smaller plasmids remained unidentified. Interestingly, we observed that the genomes of strains W6G and W7G were almost identical. This reflects the previous report by Keymer *et al*. In their original sampling study, the authors claimed that, based on CGH, 30 unique genotypes were identified within their collection of

bioRxiv preprint doi: https://doi.org/10.1101/2020.07.24.220178; this version posted July 24, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

192 41 environmental strains, while several genotypes were sampled multiple times from distinct
193 sampling events. Indeed, upon direct comparison of the W6G and W7G genomes, we
194 observed identities of 99.98%, 99.997% and 99.998% for chromosome 1, chromosome 2, and
195 the megaplasmid, respectively. Both strains possessed the same genes in all three replicons,
196 and the few observed differences were primarily single nucleotides indels (which might, in
197 part, reflect sequencing artifacts). These data suggest almost clonality between these two
198 environmental samples (W6G and W7G). In addition, our whole-genome sequencing data
199 showed that strains E7G and SA7G of clade D also had high levels of identity, which again
200 confirmed the previous CGH data (Keymer *et al.*, 2007). The majority of other strains
201 differed more significantly and therefore allowed us to test links between specific phenotypes
202 and the corresponding genotypes.
203
204 **Only a subset of environmental isolates block amoebal predation**
205 We initially wondered how these environmental isolates would behave when confronted with
206 predatory grazers, such as bacterivorous amoebae. We therefore tested the representative 7th
207 pandemic strain A1552 and the above-described collection of environmental strains for their
208 ability to defend themselves against the grazing soil amoeba *Dictyostelium discoideum*. As
209 shown in Figure 1A, a clade-specific behavior was observable in which strains from clade A
210 and B appeared as non-toxic to *D. discoideum* and were efficiently grazed on while clade C
211 and D strains completely resisted amoebal grazing. We concluded that anti-amoebal defenses
212 significantly vary among the different environmental *V. cholerae* isolates.
213
214 **Enhanced hemolysin production in environmental *V. cholerae* does not contribute to**
215 **their anti-amoebal behavior**

bioRxiv preprint doi: https://doi.org/10.1101/2020.07.24.220178; this version posted July 24, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

As we observed that a subset of the environmental strains had a strong anti-amoebal effect, we wondered whether this grazing inhibition was linked to toxic effectors of *V. cholerae*. Our group had previously demonstrated that the proper timing of the production or activity of the pore-forming toxin hemolysin HlyA was essential for pandemic *V. cholerae* to form a replication niche inside the aquatic amoebae *A. castellanii*. Indeed, while pandemic *V. cholerae* exerted a tight regulation over this toxin and thereby successfully infected the amoebae's contractile vacuole, constitutive hemolysin activity by the environmental isolates killed the host prematurely, preventing the formation of this replication niche (Van der Henst *et al.*, 2018). We speculated that hemolysin activity could also be involved in the resistance against *D. discoideum* grazing that we observed for the environmental strains from clades C and D. To follow up on this hypothesis, we first tested all environmental WT strains for hemolysis on blood agar plates. As shown in Figure 1B, all isolates had strong hemolytic activity, especially when compared with the pandemic strain A1552. To ensure that the hemolysis was indeed caused by HlyA's activity, we interrupted the *hlyA* gene (loci comparable to locus tag VCA0219 in reference strain N16961) in a subset of the environmental strains through the integration of an antibiotic resistance cassette (the wild-type (WT) versions of the environmental strains as well as their respective mutants are listed in Table S1). The selection of this subset of environmental isolates was thereby based on two criteria: i) to represent each clade; and ii) to select those strains that showed efficient chitin-induced natural transformability, which allowed for genetic manipulation of those strains. For these representative strains, we confirmed that the observed blood cell lysis was indeed caused by the specific hemolysin HlyA (Fig. 1C). After this confirmation, we tested the strains in the amoebal grazing assay. However, to our surprise, the hemolysin-deficient mutants behaved the same way as their parental WT strains, indicating that HlyA is not responsible for the strong anti-amoebal behavior that we observed for the clade C and D

bioRxiv preprint doi: https://doi.org/10.1101/2020.07.24.220178; this version posted July 24, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

241 strains (Fig. 1D). These results therefore suggest that, at least under the tested conditions,
242 HlyA does not play an important role in anti-amoebal grazing defense.
243
244 **Environmental isolates constitutively produce T6SS**
245 As HlyA activity was ruled out as a defense mechanism, we moved on to check the potential
246 involvement of other molecular weapons. We next considered the T6SS, due to its
247 widespread occurrence in Gram-negative bacteria. Notably, *V. cholerae*'s T6SS was initially
248 discovered due to its anti-eukaryotic activity that allowed the non-pandemic *V. cholerae*
249 strain V52 to avoid predation by *D. discoideum* (Pukatzki *et al.*, 2006). Indeed, previous
250 studies had indicated that, in contrast to the 7th pandemic strains, non-pandemic toxigenic *V.*
251 *cholerae* (such as the two O37 serogroup strains V52 and ATCC25872; Pukatzki *et al.*, 2006;
252 Basler *et al.*, 2012; Van der Henst *et al.*, 2018) as well as environmental isolates (Unterweger
253 *et al.*, 2012; Bernardy *et al.*, 2016; Crisan and Hammer, 2020) maintain constitutive T6SS
254 activity. This is in contrast to the silenced T6SS of pandemic strains, which are primarily
255 induced by chitin or low c-di-GMP levels, concomitantly with natural competence and the
256 production of extracellular enzymes, respectively (Borgeaud *et al.*, 2015; Watve *et al.*, 2015;
257 Metzger and Blokesch, 2016; Metzger *et al.*, 2016; Veening and Blokesch, 2017). To check
258 whether the T6SS protects clade C and D strains, we first tested the general T6SS activity of
259 the environmental isolates. As shown in Figure 2A, we observed that the environmental
260 strains efficiently eradicated *Escherichia coli* prey bacteria. Only strain SA3G of clade B
261 reproducibly killed prey with a reduced efficiency, even though residual T6SS activity was
262 still observed when compared with the nonkilling pandemic strain A1552 (Fig. 2A).
263 However, these data alone do not unambiguously show whether the observed prey
264 effacement was indeed T6SS-dependent or was instead the result of any other modes of
265 interbacterial competition, such as contact-dependent inhibition, toxin secretion, bacteriocins,

bioRxiv preprint doi: https://doi.org/10.1101/2020.07.24.220178; this version posted July 24, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

etc. (Hibbing *et al.*, 2010; Stubbendieck and Straight, 2016; Granato *et al.*, 2019). We therefore deleted the T6SS sheath protein-encoding gene *vipA* in each of the clade-representing strains and confirmed their lost T6SS activity by scoring for Hcp secretion. As illustrated in Figure 2B, all strains were able to produce Hcp protein but only the T6SS-active parental environmental strains were able to also secrete this protein into the supernatant. These data are therefore in agreement with the idea that the strains' T6SS is indeed constitutively active (at least under the tested conditions) and is the reason behind the observed interbacterial killing phenotype (Fig. 2A). We confirmed the latter idea by comparing the WT and *vipA*-minus derivatives' killing ability in an interbacterial competition assay using *E. coli* as prey (Fig. 2C).

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

Having recognized that all environmental strains constitutively produce their T6SS, we moved on to assess the involvement of this machinery in the anti-amoebal defense of clade C and D strains. Indeed, the observed inhibition of amoebal plaque formation by a subset of the environmental isolates was consistent with previous work by Unterweger and colleagues (Unterweger *et al.*, 2012). These authors had studied four environmental isolates from estuaries of the Rio Grande delta for anti-amoebal and anti-bacterial activity and observed that two of these isolates could not resist amoebal predation. These strains were, however, also unable to kill *E. coli* prey, and the reason for this interbacterial noncompetitiveness was a frameshift mutation in the intermediate T6SS regulatory protein-encoding gene *vasH* (Pukatzki *et al.*, 2006; Unterweger *et al.*, 2012). In contrast, apart from one exception (strain SA3G), all of the tested environmental isolates in our study efficiently eradicated *E. coli* prey (Fig. 2A), indicating that the T6SS was, in general, functional and active. We therefore reassessed the amoebal plaque formation against the genetically modified T6SS mutant

bioRxiv preprint doi: https://doi.org/10.1101/2020.07.24.220178; this version posted July 24, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

291 strains. As shown in Figure 2D, clade C and D isolates, whose parental WT strains
292 completely blocked amoebal predation, became nontoxic when their T6SS was inactivated,
293 indicating that their anti-amoebal defense was indeed linked to the T6SS and uniquely caused
294 by the latter.

295 Considering that all of the environmental isolates have a constitutively active T6SS
296 under the tested conditions and efficiently killed other bacteria (Fig. 2A), we wondered why
297 only the strains from clades C and D were able to use their T6SS as an anti-eukaryotic
298 defense tool. To answer this question and to also characterize the full E/I modules of these
299 strains, we inspected the T6SS clusters in the new genomic sequencing data and observed a
300 clear clade specificity. Only those strains belonging to clades C and D encoded evolved
301 VgrG1 proteins with a C-terminal actin cross-linking domain (ACD) (Figs. 2 and 3; Table
302 S3). Strains from clades A and B, on the other hand, encoded only structurally relevant
303 VgrG1 proteins without an evolved effector domain. Importantly, pandemic strains also
304 encoded such an ACD as part of *vgrG1*, but, as noted above, these strains do not produce
305 functional T6SSs without specific environmental cues and therefore show neither anti-
306 bacterial (Fig. 2A) nor anti-amoebal behavior (Fig. 2D) under the tested conditions.

307 The ACD of VgrG1 in non-pandemic but toxigenic O37 serogroup strain V52 (Chun
308 *et al.*, 2009), which produces its T6SS constitutively, was previously shown to be involved in
309 *V. cholerae*'s toxicity towards *D. discoideum* and macrophages (Pukatzki *et al.*, 2006, 2007;
310 Ma *et al.*, 2009). Moreover, this VgrG1-ACD was also responsible for intestinal
311 inflammation and cholera toxin-independent fluid accumulation in an infant mouse model of
312 infection (Ma and Mekalanos, 2010). Furthermore, the VgrG1-ACD of the 7th pandemic *V.*
313 *cholerae* strain C6706 was implicated in alternating intestinal peristalsis of zebrafish larvae,
314 leading to the expulsion of preinoculated commensal bacteria (Logan *et al.*, 2018). However,
315 this effect was only observable upon constitutive T6SS expression using a genetically

bioRxiv preprint doi: https://doi.org/10.1101/2020.07.24.220178; this version posted July 24, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

316 engineered derivative of this pandemic strain in which T6SS production occurred based on
317 artificial expression of the gene encoding the quorum sensing- and chitin-linked transcription
318 factor QstR (Lo Scrudato and Blokesch, 2013; Borgeaud *et al.*, 2015; Watve *et al.*, 2015;
319 Jaskólska *et al.*, 2018; Logan *et al.*, 2018).

320 Given this previous work on toxigenic strains, we tested whether the ACD of the clade
321 C and D environmental isolates was likewise causative of the observed anti-amoebal
322 response. To do so, we first generated truncated versions of VgrG1 that lacked the evolved
323 ACD domain-containing C-terminus (Fig. 2E). Importantly, these vgrG1ΔACD strains
324 maintained their full anti-bacterial competitiveness (Fig. 2F), which indicates that the ACD
325 deletion did not impact the general assembly and/or activity of the T6SS machinery.
326 However, as shown in Figure 2G, the amoebal grazing ability was restored on the lawns
327 formed by the two vgrG1ΔACD environmental *V. cholerae* strains and resulted in equal (e.g.,
328 for strain SL6Y) or intermediate (e.g., strain SA10G) amoebal plaque numbers compared
329 with those numbers observed for complete T6SS-defective mutants. The non-pandemic but
330 toxigenic control strain ATCC25872 (Table S1; Aldova *et al.*, 1968), which, like strain V52,
331 is constitutive T6SS active (Van der Henst *et al.*, 2018), also lost its anti-amoebal activity in
332 the absence of the T6SS core structure or when the ACD of VgrG1 was missing (Fig. 2G).
333 We therefore concluded that some environmental *V. cholerae* might have evolved ACD-
334 extended VgrG-encoding genes, as the ACD's anti-eukaryotic activity protects them from
335 environmental grazers. Whether pandemic strains subsequently horizontally acquired the
336 ACD encoding region as previously suggested (Kirchberger *et al.*, 2017), or whether the last
337 common ancestor between the pandemic and environmental lineages already contained this
338 specialty, that was later then lost from some wild strains, is currently unclear.

339

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

bioRxiv preprint doi: https://doi.org/10.1101/2020.07.24.220178; this version posted July 24, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

As briefly mentioned in the previous section, the ACD-minus mutant of the clade D strain SA10G showed reduced but still significant residual T6SS-dependent anti-amoebal activity. We therefore wondered whether this could be explained by the presence of another anti-eukaryotic effector in one or several of the T6SS clusters. Furthermore, we were also interested in characterizing the full set of E/I modules in these strains, which would allow us to speculate about the strains' competitive potential against one another. Using the newly assembled genomic data together with the previously reported genome sequence of clade B strain SA5Y (Matthey *et al.*, 2018; Matthey *et al.*, 2019) we determined the E/I modules of these 15 environmental strains and predicted their function based on BlastP analyses (Fig. 3 and Table S3). Moreover, to arrange the E/I pairs into putative compatible groups, we defined their modules based on a percentage amino-acid identity of at least 30%, which is a typing approach previously applied (Unterweger *et al.*, 2014; Kirchberger *et al.*, 2017). Interestingly, we found evidence for several orphan immunity genes, meaning immunity genes that no longer coexisted with and were adjacent to a cognate effector-encoding gene, as had been previously reported by Kirchberger and colleagues (Kirchberger *et al.*, 2017). These orphan genes were located in the 3' regions of the T6SS clusters and were likewise classified according to their module type (Fig. 3 and Table S3). The characterization of the E/I modules and orphan immunity loci was restricted to the well-characterized T6SS clusters of *V. cholerae* (e.g., the large cluster as well as auxiliary clusters 1 and 2), as the reported auxiliary cluster 3 (E/I pair TseH/TseI; Altindis *et al.*, 2015) was absent from all environmental isolates. This finding is consistent with a preprinted study showing that this auxiliary cluster represents a pandemic strain-associated mobile genetic element (Santoriello *et al.*, 2019) (Fig. 3 and Table S3). Moreover, the recently characterized auxiliary clusters 4 and 5 were also absent from the environmental isolates (Labbate *et al.*, 2016; Crisan *et al.*, 2019). Instead, automatic annotations linked to a manual analysis identified the presence of a novel

bioRxiv preprint doi: https://doi.org/10.1101/2020.07.24.220178; this version posted July 24, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

T6SS cluster in strain SP7G (clade C), which we defined as auxiliary cluster 6 (Fig. 3, Fig. S1, and Table S3). This cluster resides in the small chromosome 2 of strain SP7G and has the same gene order as auxiliary clusters 1 and 2, namely: *hcp*, *vgrG*, a putative adaptor-protein encoding gene (coding for a DUF4123 domain as described for Tap1 and VasW; Liang *et al.*, 2015; Unterweger *et al.*, 2015) plus two genes encoding proteins of unknown function, which we speculate represent an E/I pair. However, BlastP analysis identified no significant homologous proteins. Interestingly, our search for aux 6 cluster in the PATRIC nucleotide sequence database (Wattam *et al.*, 2017) suggested that this cluster is prominently represented in several *V. cholerae* strains, including 2013 environmental isolates from Bangladesh and an O35 strain (1311-69) isolated in 1969 from a patient in India (Bishop-Lilly *et al.*, 2014).

In the large T6SS cluster, pandemic *V. cholerae* strains, such as A1552 carry an A-type E/I module (Fig. 3 and Table S3), for which, in fact, the peptidoglycan cell wall degradation effector corresponds to the C-terminal domain of the evolved VgrG3 protein (Zheng *et al.*, 2011; Brooks *et al.*, 2013). When analyzing this locus in the environmental *V. cholerae* isolates, we noticed that the evolved nature of VgrG3 was conserved among these strains, except for clade C strain SP7G in which VgrG3 is solely a structural T6SS component that is followed by a B-type E/I module. How this effector is attached to the VgrG tip protein is, however, unclear, as no adaptor protein such as those encoded by *tap1* and *vasW* could be identified. This B-type effector is predicted to have a cellular adhesion function (Unterweger *et al.*, 2014). Interestingly, strains W10G (clade A) and SA3G (clade B) contain a pandemic-like A-type E/I module (Fig. 3), while the other strains carried a wide variety of E/I pairs, as described in detail in Table S3. Moreover, even though the amino acid identity of these effectors is below 30% when comparing different types, most of them have a common predicted function, namely, peptidoglycan degradation (Table S3). Therefore, with

bioRxiv preprint doi: https://doi.org/10.1101/2020.07.24.220178; this version posted July 24, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

391 the exception of strain SP7G, all the environmental isolates have a dedicated anti-bacterial
392 E/I module in the T6SS large cluster (Fig. 3 and Table S3). Interestingly, pairwise
393 comparisons of the effector and immunity protein sequences among different strains showed
394 that strains harboring E/I modules from the same family had 100% immunity identity (Table
395 S4). Given the diversity of E/I modules in this cluster when comparing all strains, the
396 complete identity of immunity proteins from the same family could indicate recent
397 acquisition by horizontal gene transfer. Furthermore, L6G and SL6Y (clade C) are the only
398 strains that also harbor one or several orphan immunity loci after the E/I module in this large
399 T6SS cluster (Fig. 3).

400 These orphan immunity loci are distinct in their type when compared with the current
401 resident E/I modules. When we searched the PATRIC translated nucleotide sequence
402 database (Wattam *et al.*, 2017) using the protein sequence from these orphan loci as the
403 query, we noticed that these genes are only found as orphan loci in other genomes. The only
404 exception was the second orphan locus in strain SL6Y, which encodes a G-type immunity
405 protein (marked by "#" in Fig. 3). The predicted protein showed 98.4% identity to true
406 immunity proteins (e.g., those encoded directly adjacent to an effector gene) from several *V.*
407 *cholerae* strains. Among those was strain 2633-78, an O1 CTX-negative isolate collected
408 from sewage in Brazil in 1978. Interestingly, this strain was experimentally tested in a
409 previous study, where it was shown to have an active T6SS (Bernardy *et al.*, 2016).

410 The auxiliary cluster 1 contains the structural or evolved (e.g., encoding C-terminal
411 ACD) *vgrG1* gene, as mentioned above (Figs. 2 and 3). In addition, in the case of pandemic
412 *V. cholerae*, this cluster harbors an A-type E/I module encoding the lipase effector TseL with
413 anti-bacterial and anti-eukaryotic activity (Zheng *et al.*, 2011; Dong *et al.*, 2013; Russell *et*
414 *al.*, 2013) followed by a C-type orphan immunity gene (Kirchberger *et al.*, 2017).
415 Interestingly, all of the 15 environmental strains harbor C-type E/I modules as part of this

bioRxiv preprint doi: https://doi.org/10.1101/2020.07.24.220178; this version posted July 24, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

auxiliary cluster 1, even though there is considerable polymorphism in the effector and immunity proteins (Fig. 3 and Table S5). C-type effectors have a predicted alpha/beta hydrolase domain (DUF2235), which has been previously associated with T6SS phospholipases from *E. coli* and *Pseudomonas aeruginosa* (Russell *et al.*, 2013; Flaugnatti *et al.*, 2016; Crisan *et al.*, 2019). Furthermore, with the exception of strains W10G (clade A) and SO5Y (clade B), all environmental isolates also carry between two and five orphan immunity loci downstream of the E/I pair (Fig. 3). Considering the wide variety of E/I types in the large cluster as well as in auxiliary cluster 2 (see below), the apparent conservation of C-type E/I modules in the auxiliary cluster 1 of these strains, in addition to the many C-type orphan loci (as also observed in the *in silico* study by Kirchberger *et al*. 2017) is quite intriguing and supports the idea that the C-type orphan immunity gene in the pandemic strains might be a remnant of a previous C-type E/I module of the strains' progenitor.

In pandemic *V. cholerae*, the auxiliary cluster 2 carries an A-type E/I module where the VasX effector is a pore-forming toxin (Miyata *et al.*, 2011; Miyata *et* al., 2013; Russell *et al.*, 2014). Our comparative genomic analyses showed that several of the environmental isolates likewise encode VasX-like effectors (A-type) at the same locus (Fig. 3 and Tables S3 and S6), while other strains carry D- and E-type E/I modules. While D-type effectors have been predicted to foster peptidoglycan degradation, E-type effectors are predicted to form pores, like VasX (Unterweger *et al.*, 2014). We therefore reasoned that the residual ACD-independent T6SS-dependent anti-amoebal impact observed for strain SA10G (clade D) might be caused by this E-type effector from auxiliary cluster 2, especially as the latter is missing from clade C strain SL6Y in which removal of the ACD from VgrG1 was almost equivalent to a complete T6SS inactivation (Fig. 2G). Interestingly, our analyses suggest that these E-type effectors also contain a common peptidoglycan-binding domain (e.g., an N-terminal Lysin Motif; Buist *et al.*, 2008), which might render them bifunctional against

bioRxiv preprint doi: https://doi.org/10.1101/2020.07.24.220178; this version posted July 24, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

bacteria and eukaryotes (Fig. 3 and Table S3). Future work will address the exact characteristics of the putative E/I modules in more detail.

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

Previous studies showed that a plethora of T6SS-transported effectors have active anti-bacterial purposes and that cognate immunity proteins are required to protect the producer or its siblings from intoxication (Dong *et al.*, 2013; Durand *et al.*, 2014; Russell *et al.*, 2014; Unterweger *et al.*, 2014). Strains with matching E/I modules could therefore coexist in the same environment (Unterweger *et al.*, 2014), while competitive strains might clear a niche and propagate inside this niche in a clonal manner (McNally *et al.*, 2017; Speare *et al.*, 2018). Importantly, as these E/I modules seem to move horizontally in an as yet uncharacterized manner (Unterweger *et al.*, 2014; Salomon *et al.*, 2015), the level of compatibility between strains will not follow their phylogenetic relatedness. We therefore sought to experimentally probe the strains' compatibility or competitiveness within this collection of environmental isolates. Indeed, while we scored T6SS activity against a laboratory strain of *E. coli* (see Fig. 2), we wondered how the *V. cholerae* strains would behave when exposed to each other. Consistent with the fact that the E/I modules carried by the environmental strains are considerably different from those of pandemic *V. cholerae* (Fig. 3 and Table S3), we experimentally demonstrated that these bacteria efficiently outcompeted the T6SS-silent pandemic strain A1552 (Fig. 4A). Next, we tested interbacterial competition between the clade-representative strains in an assay in which the prey strains had their T6SS inactivated. As shown in Figure 4B, reciprocal killing of the T6SS-positive parental strains occurred and only kin strains were immune to the toxic assaults. However, as the E/I module types seemed more similar within clades than across clades, we extended our analyses and tested all

bioRxiv preprint doi: https://doi.org/10.1101/2020.07.24.220178; this version posted July 24, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

pairwise combinations (this time, with both partners T6SS positive). The underlying rationale was that previous studies had identified different types of E/I modules *in silico* (Unterweger *et al.*, 2014; Kirchberger *et al.*, 2017; Crisan *et al.*, 2019) but most pairwise competition assays were primarily linked to patient isolates and not to a larger collection of environmental isolates. Indeed, as described above and shown in Figure 3, our *in silico* prediction showed that strains within clades often contained similar E/I module types (Tables S4 to S6), even though the amino acid identity threshold for this categorization is rather low (30% as previously defined; Unterweger *et al.*, 2014; Kirchberger *et al.*, 2017).

As shown in Figure 4, we observed interesting and complex phenotypes in these extensive pairwise killing experiments. For instance, strains W6G and W7G from clade C were fully protected against each other's attacks (Fig. 4 panels M and N). This finding is consistent with the 100% identity of all three T6SS clusters, including the orphan genes from aux cluster 1 (Tables S4 to S6), and their overall near clonality as described above. Clade D strains have the same E/I module types in all T6SS clusters. However, upon closer inspection of their E/I protein sequences, we can observe a 100% sequence identity among the three strains only in the large and aux 1 cluster-encoded proteins. For the aux 2 cluster, the encoded E/I proteins from strains E7G and SA7G are 100% identical, while the E/I pair carried by strain SA10G has only 95%/73% identity when compared with the other two strains' E/I pair (Tables S4 to S6). Consequently, and as shown in Figure 4 (panel O to Q), strains E7G and SA7G are fully compatible with each other, while strain SA10G can kill and be killed by the other two strains. We therefore concluded that the identity level in only one of the T6SS cluster-encoded immunity proteins causes the competitiveness among these strains.

Surprisingly, clade B strains SA5Y and SL4G have 100% identity in all three T6SS cluster-encoded immunity proteins (Tables S4 to S6) but are still able to kill each other with

bioRxiv preprint doi: https://doi.org/10.1101/2020.07.24.220178; this version posted July 24, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

491 considerable efficiency (Fig. 4, panels E and F). Why this is the case is currently unclear.
492 However, we speculate that expression or immunity protein production might be impaired in
493 those strains or that additional T6SS E/I modules are hidden in the strains' genomes, which
494 were not easily identifiable based on a lack of the hallmark genes *paar*, *hcp* and *vgrG* in their
495 vicinity.

496 Clade A strain W10G carries pandemic-like A-type E/I modules in the large and aux2
497 clusters (Fig. 3). Interestingly, this strain can kill many of the other environmental strains
498 remarkably well, such as clade B strains SA5Y, SL4G and SL5Y, and clade C strains SL6Y
499 and SP6G. Notably, the T6SS active toxigenic strain V52 (which harbors A-type E/I modules
500 in all T6SS clusters) was previously shown to outcompete strains carrying different
501 combinations of E/I modules (Unterweger *et al.*, 2014). Indeed, the AAA (A-type in all
502 clusters) E/I modules is conserved in pandemic *V. cholerae*, even in those strains that caused
503 former pandemics (e.g., 6th pandemic O1 classical strains), as well as in non-pandemic but
504 toxigenic isolates. It was therefore speculated that this combination might be advantageous in
505 a disease context (Unterweger *et al.*, 2014). Clade B strain SA3G also contains A-type E/I
506 modules in the large and aux 2 clusters, but it does not kill other strains at the same level as
507 strain W10G. However, as noted above, this strain is even less efficient against laboratory *E.*
508 *coli* prey strains (Fig. 2A), which could mask its full effector toxicity potential.

509 Finally, a very interesting pairwise comparison is that of clade C strains SP7G and
510 L6G. While L6G as a prey is very efficiently eliminated by strain SP7G (Fig. 4, panel I),
511 SP7G prey is resistant to L6G intoxication (Fig. 4, panel L). These strains contain E/I sets of
512 different families in the large cluster and show 78% and 99.7% identity in the immunity
513 proteins encoded in aux cluster 1 and 2, respectively. Based on these differences, we would
514 expect that these strains fully compete with each other; however, killing only occurs with
515 SP7G as the predator. We therefore speculate that the large cluster-encoded K-type effector

bioRxiv preprint doi: https://doi.org/10.1101/2020.07.24.220178; this version posted July 24, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

516 domains of strain L6G's evolved VgrG3 as well as SP7G's noncanonical structural VgrG3
517 and its adjacently encoded putative effector protein might not be functional peptidoglycan
518 destruction enzymes and therefore not necessarily active in interbacterial competition. Why a
519 low level of protein identity in the aux cluster 2-encoded immunity protein is sufficient to
520 protect strain SP7G from L6G assaults is currently unclear. However, it is tempting to
521 speculate that the *vice versa* interaction leads to L6G killing due to the additional aux cluster
522 6 that is carried by strain SP7G. Further work is therefore necessary in order to delve deeper
523 into these observed phenotypes.

524 Overall, closer inspection of all of the pairwise killing data attests to the complexity
525 behind the T6SS compatibility code. As mentioned previously, even though some strains
526 might harbor the same E/I families in the T6SS clusters, pairwise comparisons of these
527 proteins shows that quite frequently their identity is not 100% (Tables S4 to S6), which
528 appears to be necessary to allow coexistence (also observed in Speare *et al.*, 2018). Our data
529 therefore support what Unterweger and colleagues (Unterweger *et al.*, 2014) initially
530 speculated, namely, that the compatibility between strains seems to follow the level of
531 polymorphism of their immunity proteins and that diversity in only one cluster-encoded
532 protein is sufficient to drive competition.

533

534 **Conclusion**

535 In this study, we investigated two minor virulence factors, the pore forming hemolysin and
536 the T6SS, in a set of fifteen environmental *V. cholerae* strains. We assessed the relevance of
537 these molecular weapons as defense mechanisms against amoebal predation (hemolysin and
538 T6SS) and in the context of bacterial warfare (T6SS only). We showed that all of these
539 environmental isolates possess a constitutively active T6SS and are able to use the machinery
540 as a bacterial killing device. In contrast, only a subset of these strains was able to efficiently

bioRxiv preprint doi: https://doi.org/10.1101/2020.07.24.220178; this version posted July 24, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

suppress grazing by *D. discoideum* amoebae, a phenotype that was dependent on the eukaryote-specific ACD of the evolved VgrG1 T6SS effector. Careful *in silico* identification unveiled an extensive T6SS repertoire of E/I pairs and orphan immunity loci. Consistent with this finding, we observed extensive interbacterial competition under pairwise coculture conditions whereby mutual compatibility was rarely achieved. Importantly, our study also confirmed that both molecular weapons, the T6SS and the hemolysin toxin, are constitutively active in the environmental isolates in contrast to their tight regulation in the well-studied pandemic patient isolates. Future work is therefore required to decipher how this differential production pattern is achieved in pandemic versus non-pandemic strains and whether this tight regulatory control might provide specific benefits to the former strains.

Experimental procedures

Bacterial strains and growth conditions

The bacterial strains (*V. cholerae*, *E. coli* and *K. pneumoniae*) used in this study are listed in Table S1. Unless otherwise stated, all strains were grown aerobically in Lysogeny broth (LB; 10 g/L of tryptone, 5 g/L of yeast extract, 10 g/L of sodium chloride; Carl Roth) or on LB agar plates at 30°C. Half-concentrated defined artificial seawater medium (0.5×DASW) containing HEPES and vitamins (Meibom *et al.*, 2005) was used for growth on chitinous surfaces for strain construction based on chitin-induced natural transformation (see below).

D. discoideum amoebae (strain Ax2 Ka) were cultured in HL5 medium supplemented with glucose (Formedium, UK). For amoebal grazing assays (e.g., plaque formation assays; see below), SM/5 medium (final concentrations: 2 g/L of glucose, 2 g/L of bacto peptone, 2 g/L of yeast extract, 0.2 g/L of $MgSO_4$ $7H_2O$, 1.9 g/L of KH_2PO_4, 1 g/L of K_2HPO_4; pH 6.4), was mixed with 2% agar to prepare SM/5 plates (20 mL/plate) (Sussman, 1987). $CaCl_2$ (50 µM)-supplemented Sörensen's buffer (8 g /4 L of KH_2PO_4, 1.16 g /4 L of Na_2HPO_4; pH 6;

bioRxiv preprint doi: https://doi.org/10.1101/2020.07.24.220178; this version posted July 24, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

566 Gerisch *et al*., 1967) was used as the washing and resuspension buffer for the amoebae and
567 the bacteria that were used in the plaque formation assay.

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

571

572 **Genetic engineering**

573 *V. cholerae* strains were genetically modified using chitin-induced transformation as
574 previously described (Marvig 2010; Silva and Blokesch, 2010; Blokesch, 2012; Borgeaud
575 and Blokesch, 2013). This method relies on natural transformation triggered by growth on
576 chitin followed by the addition of a PCR fragment that carried the desired genetic change. To
577 achieve higher numbers of transformants, the protocol was slightly modified. The PCR
578 fragments were added twice (24h and 36h after bacterial inoculation on the chitin flakes) and
579 cells were enriched in 2×YT medium (Carl Roth) before selective plating. PCR
580 amplifications were conducted using Pwo (Roche) and GoTaq (Promega) polymerases
581 according to the suppliers' recommendations. Following initial screening by PCR (using
582 bacterial cells as the templates), genetically engineered loci were verified by Sanger
583 sequencing (Microsynth, Switzerland).

584 The rifampicin-sensitive *V. cholerae* strain A1552-Rif^S was generated by a
585 combination of natural cotransformation (Dalia *et al.*, 2014) and our previously described
586 counterselectable Trans2 approach (Van der Henst *et al.*, 2018). To this end, a 4kb PCR
587 fragment was amplified, harboring a mutation, which restored the native *rpoB*-encoded
588 protein (F531S substitution) from its mutated version in the parental strain A1552
589 (RpoB[S531F]; Matthey *et al.*, 2018). *V. cholerae* A1552 was then cotransformed with this
590 fragment and another 3,924 bp fragment containing flanking regions matching *lacZ* and two

bioRxiv preprint doi: https://doi.org/10.1101/2020.07.24.220178; this version posted July 24, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

591 selection markers (*aph* and *pheS**; Table S1). Transformants were selected on kanamycin-
592 containing agar plates and the lost rifampicin resistance was scored based on replica plating
593 on plates +/- rifampicin using a velvet cloth. A second round of natural transformation
594 followed to restore the *lacZ* gene by adding a WT *lacZ* PCR fragment to chitin-grown cells
595 followed by a counter selection of the *pheS** allele on 4-chloro-phenylalanine (20mM)-
596 containing agar plates as previously described (Van der Henst *et al.*, 2018). To confirm the
597 restoration of native *rpoB*, genomic DNA of A1552 Rif^S was isolated and the PCR-amplified
598 *rpoB* gene was Sanger-sequenced.
599
600 **Amoebal grazing assay**
601 To determine the predatory capacity of *D. discoideum* on bacterial lawns of *V. cholerae*,
602 plaque formation was scored following a previously described protocol (Pukatzki *et al.*,
603 2006) with minor modifications. Briefly, bacteria were cultured overnight in LB medium at
604 30°C and harvested by centrifugation. The cell pellet was washed and resuspended in SorC
605 buffer (Gerisch *et al.*, 1967), and then diluted with SorC to reach an optical density at 600 nm
606 (OD_{600}) of 5.5 in a final volume of 400 µL. Cultured *D. discoideum* amoebae (in HL5 with
607 glucose; Formedium, UK) were detached from culture dishes using cell scraper (SPL Life
608 Sciences) and collected by centrifugation (3 min 1,000 rcf), resuspended in SorC buffer, and
609 enumerated in a KOVA counting chamber (KOVA International, USA). The amoebal
610 concentration was adjusted to 2×10^4 cells/ mL and 20 µL of this suspension (corresponding
611 to ~400 amoebal cells) was mixed with the 400 µL of bacterial suspension. The mixture was
612 gently spread on two parallel SM/5 plates using a plastic rake (VWR), resulting in technical
613 replicates. The plates were wrapped in aluminum foil and incubated at 24°C for 5 days. After
614 this incubation period, *D. discoideum* plaque numbers were enumerated. As a positive
615 control, we included a frequently used nonencapsulated *Klebsiella* strain (Benghezal *et al.*,

bioRxiv preprint doi: https://doi.org/10.1101/2020.07.24.220178; this version posted July 24, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

2006) for which the resulting plaque numbers were set to 100%. Three biologically independent experiments were performed. The individual experimental data points (mean of technical replicates) as well as the overall average of the independent experiments (+/- standard deviation) are shown in each graph. A two-tailed Student's *t*-test was performed to determine statistical significance.

Hemolysin activity

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

Interbacterial killing assays

Bacterial killing was assessed following a previously established assay with minor modifications (Borgeaud *et al.*, 2015). The prey cells (*E. coli* or *V. cholerae*, as indicated) and the respective predator bacteria were mixed at a ratio of 1:10 and spotted onto paper filters on prewarmed LB agar plates. After 4 h of incubation at 37°C, the bacteria were resuspended, serially diluted, and spotted onto antibiotic-containing (rifampicin or streptomycin) LB agar plates to enumerate the colony-forming units (shown as CFU/ml). The majority of these killing experiments were performed using exponentially growing *V. cholerae* (OD_{600} ~1). For the pairwise killing experiments of all environmental strains, OD_{600}-adjusted overnight cultures were used, which resulted in biologically similar outcomes to the samples derived from exponentially growing cultures. Statistically significant differences were determined on log-transformed data (Keene, 1995) by a two-tailed Student's

bioRxiv preprint doi: https://doi.org/10.1101/2020.07.24.220178; this version posted July 24, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

t-test of three biologically independent replicates. If no prey bacteria were recovered, the value was set to the detection limit to allow for statistical analysis.

SDS-PAGE and western blotting

To check the production of the Hcp protein, cell lysates were prepared as described previously (Metzger *et al.*, 2016). In brief, exponentially growing bacteria (~3 hours of growth after a 1:100 back dilution from overnight cultures) were pelleted and then resuspended in Laemmli buffer, adjusting for the total number of bacteria according to the cultures' OD_{600} values. To check for T6SS-secreted Hcp, 1.5 ml of the culture supernatant was filter sterilized (0.2-µm filter; VWR) and the proteins were precipitated using trichloroacetic acid (TCA). The precipitated proteins were washed with acetone before being resuspended in 30 µL of Laemmli buffer. All samples were heated at 95°C for 15 min.

Proteins were separated by sodium dodecyl sulfate (SDS)-polyacrylamide gel electrophoresis (PAGE) using 15% gels and then western blotted as previously described (Lo Scrudato and Blokesch, 2012). Primary antibodies against Hcp (Eurogentec; Metzger *et al.*, 2016) were used at 1:5,000 dilution while anti-Sigma70-HRP antibodies (BioLegend, USA distributed via Brunschwig, Switzerland) were diluted 1:10,000 and served as a loading control. Goat anti-rabbit horseradish peroxidase (HRP) (diluted at 1:20,000; Sigma-Aldrich, Switzerland) was used as the secondary antibody against the anti-Hcp primary antibody. Lumi-Light^{PLUS} western blotting substrate (Roche, Switzerland) served as the HRP substrate. The signals were detected using a ChemiDoc XRS+ station (BioRad).

Preparation of genomic DNA for whole-genome sequencing

Genomic DNA (gDNA) was purified from 2 ml of an overnight culture of the respective strain. DNA extraction was performed using 100/G Genomic-tips together with a Genomic

bioRxiv preprint doi: https://doi.org/10.1101/2020.07.24.220178; this version posted July 24, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

665 DNA buffer set as described in the manufacturer's instructions (Qiagen). After precipitation,
666 the DNA samples were washed twice with cold 70% ethanol and dissolved in Tris buffer (10
667 mM Tris-HCl, pH 8.0).
668
669 **Long-read PacBio genome sequencing**
670 Sample preparation and genome sequencing was performed by the Genomic Technology
671 Facility of the University of Lausanne (Switzerland) using standard protocols. Briefly, DNA
672 samples were sheared in Covaris g-TUBEs to obtain fragments with a mean length of 20 kb.
673 The sheared DNA was used to prepare each library with the PacBio SMRTbell template prep
674 kit 1 (Pacific Biosciences) according to the manufacturer's recommendations. The resulting
675 library was size selected on a BluePippin system (Sage Science, Inc.) for molecules larger
676 than 15 kb, which excluded smaller plasmids. Each library was sequenced on one single-
677 molecule real-time (SMRT) cell with P6/C4 chemistry and MagBeads on a PacBio RS II
678 system at a movie length of 360 min. Genome assembly was performed using the protocol
679 RS_HGAP_Assembly.3 in SMRT Pipe 2.3.0, and circularization of the genomes was
680 achieved using the Minimus assembler of the AMOS software package 3.1.0 using default
681 parameters (Sommer *et al.*, 2007). The assembled genomes were initially annotated using
682 Prokka 1.12 (Seemann, 2014) but due to several incompatibilities with the NCBI database,
683 they were reannotated with their own pipeline (PGAP annotation) during NCBI submission.
684 The genomic data and NCBI accession numbers are summarized in Table S2.
685 Sequencing artifacts were observed for two genomic regions of strain W10G (clade
686 A), leading to frameshifted putative effector genes. Briefly, for aux cluster 1, the effector
687 gene was split into two ORFs (locus tags VC-W10G_01483 and VC-W10G_01482 in
688 CP053794), due to the insertion of a C in position 321 of the first ORF, which led to a
689 frameshift and consequently to an early stop codon at position 351. Similarly, in aux cluster

bioRxiv preprint doi: https://doi.org/10.1101/2020.07.24.220178; this version posted July 24, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

2, the putative effector gene was also split into two ORFs (locus tags W10G_02844 and W10G_02845 in CP053795) due to the insertion of a C in position 890, which likewise caused a frameshift followed by early termination. Both of these regions were Sanger sequenced after PCR-amplification using the same genomic DNA samples as templates that were initially used for PacBio library preparation, which confirmed that the additional C bases were in both cases a sequencing artifact and the genes were properly maintained in strain W10G (as indicated in Fig. 3).

Characterization of E/I modules from environmental strains

T6SS clusters of the environmental strains were identified by searching the Prokka-annotated genomes for conserved genes, such as *paar, vgrG* and *hcp,* as well as according to their location when aligned to the genome of the pandemic strain A1552 (Matthey *et al.*, 2018). All identified putative effectors were compared by BLAST against the NCBI database to identify conserved domains. Additional characterization was made based on the previous literature (Unterweger *et al.*, 2014; Kirchberger *et al.*, 2017). Furthermore, all E/I modules and orphan immunity loci were classified for their family type. Proteins with a sequence of less than 30% identity were considered as distinct incompatible types, as previously described (Unterweger *et al.*, 2014; Kirchberger *et al.*, 2017). For VgrG3 proteins, the typing was only based on the effector portion of the protein (Unterweger *et al.*, 2014). To determine this part of the protein, the full VgrG3 sequence from all strains was aligned. The conserved region corresponding to the VgrG part of the protein was subsequently removed, which left only the variable C-terminal effector domain. This part was then used for typing and pairwise comparisons. The putative orphan immunity proteins were queried against a translated nucleotide database (PATRIC; Wattam *et al.*, 2017) to identify homologous *bona fide* immunity proteins encoded adjacent to an effector gene in the other T6SS clusters.

bioRxiv preprint doi: https://doi.org/10.1101/2020.07.24.220178; this version posted July 24, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

Data availability

PacBio raw reads of the 14 whole-genome sequenced strains have been deposited in NCBI's Sequence Read Archive (SRA) under Bioproject accession number PRJNA633476. Details on the SRA accession numbers, BioSamples, and individual accession numbers of the *de novo* assembled and circularized genomes are provided in Table S2.

Acknowledgments

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

bioRxiv preprint doi: https://doi.org/10.1101/2020.07.24.220178; this version posted July 24, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

References

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

Ali, M., Nelson, A.R., Lopez, A.L., and Sack, D.A. (2015) Updated global burden of cholera in endemic countries. *PLoS Negl Trop Dis* **9**: 1–13.

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

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

Benghezal, M., Fauvarque, M.O., Tournebize, R., Froquet, R., Marchetti, A., Bergeret, E., *et al*. (2006) Specific host genes required for the killing of *Klebsiella* bacteria by phagocytes. *Cell Microbiol* **8**: 139–148.

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

Bishop-Lilly, K., Johnson, S., Verratti, K., Truong, L., Amy, K., Joy, A., *et al*. (2014) Genome sequencing of 15 clinical isolates, including 13 non-O1/non-O139 serogroup strains. *Genome Announc* **2**: e00893-14.

Blokesch, M. (2012) TransFLP - A method to genetically modify *Vibrio cholerae* based on natural transformation and FLP-recombination. *JoVE*: e3761.

Borgeaud, S. and Blokesch, M. (2013) Overexpression of the *tcp* gene cluster using the T7 RNA polymerase/promoter system and natural transformation-mediated genetic

bioRxiv preprint doi: https://doi.org/10.1101/2020.07.24.220178; this version posted July 24, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

763 engineering of *Vibrio cholerae*. *PLoS One* **8**: e53952.

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

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

769 Bröms, J.E., Ishikawa, T., Wai, S.N., and Sjöstedt, A. (2013) A functional VipA-VipB
770 interaction is required for the type VI secretion system activity of *Vibrio cholerae* O1
771 strain A1552. *BMC Microbiol* **13**: 96.

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

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

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

779 Chou, S., Bui, N.K., Russell, A.B., Lexa, K.W., Gardiner, T.E., LeRoux, M., *et al*. (2012)
780 Structure of a peptidoglycan amidase effector targeted to Gram-negative bacteria by the
781 type VI secretion system. *Cell Rep* **1**: 656–664.

782 Chun, J., Grim, C.J., Hasan, N.A., Lee, J.H., Choi, S.Y., Haley, B.J., *et al.* (2009)
783 Comparative genomics reveals mechanism for short-term and long-term clonal
784 transitions in pandemic *Vibrio cholerae*. *Proc Natl Acad Sci USA* **106**: 15442–15447.

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

787 Clemens, J.D., Nair, G.B., Ahmed, T., Qadri, F., and Holmgren, J. (2017) Cholera. *Lancet*

bioRxiv preprint doi: https://doi.org/10.1101/2020.07.24.220178; this version posted July 24, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

390: 1539–1549.

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

Crisan, C. V., Chande, A.T., Williams, K., Raghuram, V., Rishishwar, L., Steinbach, G., *et al.* (2019) Analysis of *Vibrio cholerae* genomes identifies new type VI secretion system gene clusters. *Genome Biol* **20**: 163.

Crisan, C. V and Hammer, B.K. (2020) The *Vibrio cholerae* type VI secretion system: toxins, regulators and consequences. *Environm Microb* (online ahead of print).

Dalia, A.B., McDonough, E.K., and Camilli, A. (2014) Multiplex genome editing by natural transformation. *Proc Natl Acad Sci USA* **111**: 8937–8942.

Deshayes, S., Daurel, C., Cattoir, V., Parienti, J.J., Quilici, M.L., and de La Blanchardière, A. (2015) Non-O1, non-O139 *Vibrio cholerae* bacteraemia: case report and literature review. *Springerplus* **4**: 575.

Domman, D., Quilici, M.-L., Dorman, M.J., Njamkepo, E., Mutreja, A., Mather, A.E., *et al.* (2017) Integrated view of *Vibrio cholerae* in the Americas. *Science* **358**: 789–793.

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

Durand, E., Cambillau, C., Cascales, E., and Journet, L. (2014) VgrG, Tae, Tle, and beyond: The versatile arsenal of type VI secretion effectors. *Trends Microbiol* **22**: 498–507.

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

Dziejman, M., Serruto, D., Tam, V.C., Sturtevant, D., Diraphat, P., Faruque, S.M., *et al.*

bioRxiv preprint doi: https://doi.org/10.1101/2020.07.24.220178; this version posted July 24, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

(2005) Genomic characterization of non-O1, non-O139 *Vibrio cholerae* reveals genes for a type III secretion system. *Proc Natl Acad Sci* **102**: 3465–3470.

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

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

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

Flaugnatti, N., Le, T.T.H., Canaan, S., Aschtgen, M.S., Nguyen, V.S., Blangy, S., *et al.* (2016) A phospholipase A1 antibacterial type VI secretion effector interacts directly with the C-terminal domain of the VgrG spike protein for delivery. *Mol Microbiol* **99**: 1099–1118.

Flaugnatti, N., Rapisarda, C., Rey, M., Beauvois, S.G., Nguyen, V.A., Canaan, S., *et al.* (2020) Structural basis for loading and inhibition of a bacterial T6SS phospholipase effector by the VgrG spike. *EMBO J* **39:** e104129.

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

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

Gerisch, G; Lüderitz, O; Ruschmann, E. (1967) Antikörper fördern die Phagozytose von

bioRxiv preprint doi: https://doi.org/10.1101/2020.07.24.220178; this version posted July 24, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

838 Bakterien durch Amöben. *Zsch Naturforsch* **22b**: 109.

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

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

843 Harris, J.B., LaRocque, R.C., Qadri, F., Ryan, E.T., and Calderwood, S.B. (2012) Cholera.
844 *Lancet* **379**: 2466–2476.

845 Hasan, N.A., Rezayat, T., Blatz, P.J., Choi, S.Y., Griffitt, K.J., Rashed, S.M., *et al.* (2015)
846 Nontoxigenic *Vibrio cholerae* non-O1/O139 isolate from a case of human gastroenteritis
847 in the U.S. Gulf Coast. *J Clin Microbiol* **53**: 9–14.

848 Heidelberg, J.F., Elsen, J.A., Nelson, W.C., Clayton, R.A., Gwinn, M.L., Dodson, R.J., *et al.*
849 (2000) DNA sequence of both chromosomes of the cholera pathogen *Vibrio cholerae*.
850 *Nature* **406**: 477–483.

851 Van der Henst, C., Vanhove, A.S., Drebes Dörr, N.C., Stutzmann, S., Stoudmann, C., Clerc,
852 S., *et al*. (2018) Molecular insights into *Vibrio cholerae*'s intra-amoebal host-pathogen
853 interactions. *Nat Commun* **9**: 3420.

854 Hersch, S.J., Watanabe, N., Stietz, M.S., Manera, K., Kamal, F., Burkinshaw, B., *et al.*
855 (2020) Envelope stress responses defend against type six secretion system attacks
856 independently of immunity proteins. *Nat Microbiol* **5:** 706–714.

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

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

861 Hood, R.D., Singh, P., Hsu, F.S., Güvener, T., Carl, M.A., Trinidad, R.R.S., *et al.* (2010) A
862 type VI secretion system of *Pseudomonas aeruginosa* targets a toxin to bacteria. *Cell*

bioRxiv preprint doi: https://doi.org/10.1101/2020.07.24.220178; this version posted July 24, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

Host Microbe **7**: 25–37.

Ichinose, Y., Yamamoto, K., Nakasone, N., Tanabe, M.J., Takeda, T., Miwatani, T., and Iwanaga, M. (1987) Enterotoxicity of El Tor-like hemolysin of non-O1 *Vibrio cholerae*. *Infect Immun* **55**: 1090–1093.

Islam, M.S., Alam, M.J., and Neogi, P.K.B. (1992) Seasonality and toxigenicity of *Vibrio cholerae* non-O1 isolated from different components of pond ecosystems of Dhaka City, Bangladesh. *World J Microbiol Biotechnol* **8**: 160–163.

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

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

Keene, O.N. (1995) The log transformation is special. *Stat Med* **14**: 811–819.

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

Kirchberger, P.C., Unterweger, D., Provenzano, D., Pukatzki, S., and Boucher, Y. (2017) Sequential displacement of type VI secretion system effector genes leads to evolution of diverse immunity gene arrays in *Vibrio cholerae*. *Sci Rep* **7**: 45133.

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

Kube, S., Kapitein, N., Zimniak, T., Herzog, F., Mogk, A., and Wendler, P. (2014) Structure of the VipA/B type VI secretion complex suggests a contraction-state-specific recycling mechanism. *Cell Rep* **8**: 20–30.

bioRxiv preprint doi: https://doi.org/10.1101/2020.07.24.220178; this version posted July 24, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

888 Labbate, M., Orata, F.D., Petty, N.K., Jayatilleke, N.D., King, W.L., Kirchberger, P.C., *et al*.
889 (2016) A genomic island in *Vibrio cholerae* with VPI-1 site-specific recombination
890 characteristics contains CRISPR-Cas and type VI secretion modules. *Sci Rep* **6**: 36891.
891 Liang, X., Moore, R., Wilton, M., Wong, M.J.Q., Lam, L., and Dong, T.G. (2015)
892 Identification of divergent type VI secretion effectors using a conserved chaperone
893 domain. *Proc Natl Acad Sci USA* **112**: 9106–9111.
894 Logan, S.L., Thomas, J., Yan, J., Baker, R.P., Shields, D.S., Xavier, J.B., *et al*. (2018) The
895 *Vibrio cholerae* type VI secretion system can modulate host intestinal mechanics to
896 displace gut bacterial symbionts. *Proc Natl Acad Sci USA* **115**: E3779–E3787.
897 Ma, A.T., McAuley, S., Pukatzki, S., and Mekalanos, J.J. (2009) Translocation of a *Vibrio*
898 *cholerae* type VI secretion effector requires bacterial endocytosis by host cells. *Cell*
899 *Host Microbe* **5**: 234–243.
900 Ma, A.T. and Mekalanos, J.J. (2010) *In vivo* actin cross-linking induced by *Vibrio cholerae*
901 type VI secretion system is associated with intestinal inflammation. *Proc Natl Acad Sci*
902 *USA* **107**: 4365–4370.
903 Matthey, N., Drebes Dörr, N.C., and Blokesch, M. (2018) Long-read-based genome
904 sequences of pandemic and environmental *Vibrio cholerae* strains. *Microbiol Resour*
905 *Announc* **7**: e01574-18.
906 McNally, L., Bernardy, E., Thomas, J., Kalziqi, A., Pentz, J., Brown, S.P., *et al*. (2017)
907 Killing by type VI secretion drives genetic phase separation and correlates with
908 increased cooperation. *Nat Commun* **8**: 14371.
909 Meibom, K.L., Blokesch, M., Dolganov, N.A., Wu, C.Y., and Schoolnik, G.K. (2005) Chitin
910 induces natural competence in *Vibrio cholerae*. *Science* **310**: 1824–1827.
911 Metzger, L.C. and Blokesch, M. (2016) Regulation of competence-mediated horizontal gene
912 transfer in the natural habitat of *Vibrio cholerae*. *Curr Opin Microbiol* **30**: 1–7.

bioRxiv preprint doi: https://doi.org/10.1101/2020.07.24.220178; this version posted July 24, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

913 Metzger, L.C., Stutzmann, S., Scrignari, T., Van der Henst, C., Matthey, N., and Blokesch,
914 M. (2016) Independent regulation of type VI secretion in *Vibrio cholerae* by TfoX and
915 TfoY. *Cell Rep* **15**: 951–958.
916 Miller, M.C., Keymer, D.P., Avelar, A., Boehm, A.B., and Schoolnik, G.K. (2007) Detection
917 and transformation of genome segments that differ within a coastal population of *Vibrio*
918 *cholerae* strains. *Appl Environ Microbiol* **73**: 3695–3704.
919 Miyata, S.T., Kitaoka, M., Brooks, T.M., McAuley, S.B., and Pukatzki, S. (2011) *Vibrio*
920 *cholerae* requires the type VI secretion system virulence factor vasx to kill
921 *Dictyostelium discoideum*. *Infect Immun* **79**: 2941–2949.
922 Miyata, S.T., Unterweger, D., Rudko, S.P., and Pukatzki, S. (2013) Dual expression profile
923 of type VI secretion system immunity genes protects pandemic *Vibrio cholerae*. *PLoS*
924 *Pathog* **9**: e1003752.
925 Mutreja, A., Kim, D.W., Thomson, N.R., Connor, T.R., Lee, J.H., Kariuki, S., *et al.* (2011)
926 Evidence for several waves of global transmission in the seventh cholera pandemic.
927 *Nature* **477**: 462–465.
928 Okada, K., Iida, T., Kita-Tsukamoto, K., and Honda, T. (2005) Vibrios commonly possess
929 two chromosomes. *J Bacteriol* **187**: 752–757.
930 Olivier, V., Haines, G.K., Tan, Y., and Fullner Satchell, K.J. (2007) Hemolysin and the
931 multifunctional autoprocessing RTX toxin are virulence factors during intestinal
932 infection of mice with *Vibrio cholerae* El Tor O1 strains. *Infect Immun* **75**: 5035–5042.
933 Onifade, T.J.M., Hutchinson, R., van Zile, K., Bodager, D., Baker, R., and Blackmore, C.
934 (2011) Toxin producing *Vibrio cholerae* O75 outbreak, United States, March to April
935 2011. *Eurosurveillance* **16**: 20–22.
936 Pukatzki, S., Ma, A.T., Revel, A.T., Sturtevant, D., and Mekalanos, J.J. (2007) Type VI
937 secretion system translocates a phage tail spike-like protein into target cells where it

bioRxiv preprint doi: https://doi.org/10.1101/2020.07.24.220178; this version posted July 24, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

938 cross-links actin. *Proc Natl Acad Sci USA* **104**: 15508–15513.

939 Pukatzki, S., Ma, A.T., Sturtevant, D., Krastins, B., Sarracino, D., Nelson, W.C., *et al*. (2006)
940 Identification of a conserved bacterial protein secretion system in *Vibrio cholerae* using
941 the *Dictyostelium* host model system. *Proc Natl Acad Sci USA* **103**: 1528–33.

942 Pukatzki, S., McAuley, S.B., and Miyata, S.T. (2009) The type VI secretion system:
943 translocation of effectors and effector-domains. *Curr Opin Microbiol* **12**: 11–17.

944 Rivera, I.N.G., Chun, J., Huq, A., Sack, R.B., and Colwell, R.R. (2001) Genotypes associated
945 with virulence in environmental isolates of *Vibrio cholerae*. *Appl Environ Microbiol* **67**:
946 2421–2429.

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

949 Russell, A.B., Leroux, M., Hathazi, K., Agnello, D.M., Ishikawa, T., Wiggins, P.A., *et al*.
950 (2013) Diverse type VI secretion phospholipases are functionally plastic antibacterial
951 effectors. *Nature* **496**: 508–512.

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

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

957 Santoriello, F.J., Michel, L., Unterweger, D., and Pukatzki, S. (2019) *Vibrio cholerae* type VI
958 secretion system auxiliary cluster 3 is a pandemic-associated mobile genetic element.
959 *bioRxiv* **doi:** https://doi.org/10.1101/868539 (preprint).

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

bioRxiv preprint doi: https://doi.org/10.1101/2020.07.24.220178; this version posted July 24, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

963 Lo Scrudato, M. and Blokesch, M. (2012) The regulatory network of natural competence and
964 transformation of Vibrio cholerae. *PLoS Genet* **8(6):** e1002778.
965 Seemann, T. (2014) Prokka: Rapid prokaryotic genome annotation. *Bioinformatics* **30**: 2068–
966 2069.
967 Shapiro, B.J., Levade, I., Kovacikova, G., Taylor, R.K., and Almagro-Moreno, S. (2016)
968 Origins of pandemic *Vibrio cholerae* from environmental gene pools. *Nat Microbiol* **2**:
969 16240.
970 Shneider, M.M., Buth, S.A., Ho, B.T., Basler, M., Mekalanos, J.J., and Leiman, P.G. (2013)
971 PAAR-repeat proteins sharpen and diversify the type VI secretion system spike. *Nature*
972 **500**: 350–353.
973 De Souza Silva, O. and Blokesch, M. (2010) Genetic manipulation of *Vibrio cholerae* by
974 combining natural transformation with FLP recombination. *Plasmid* **64**: 186–195.
975 Singh, D. V, Matte, M.H., Matte, G.R., Jiang, S., Sabeena, F., Shukla, B.N., *et al*. (2001)
976 Molecular analysis of *Vibrio cholerae* O1, O139, non-O1, and non-O139 strains: clonal
977 relationships between clinical and environmental isolates. *Appl Environ Microbiol* **67**:
978 910–921.
979 Sommer, D.D., Delcher, A.L., Salzberg, S.L., and Pop, M. (2007) Minimus: A fast,
980 lightweight genome assembler. *BMC Bioinformatics* **8**: 64.
981 Speare, L., Cecere, A.G., Guckes, K.R., Smith, S., Wollenberg, M.S., Mandel, M.J., *et al*.
982 (2018) Bacterial symbionts use a type VI secretion system to eliminate competitors in
983 their natural host. *Proc Natl Acad Sci USA* **115**: E8528–E8537.
984 Stubbendieck, R.M. and Straight, P.D. (2016) Multifaceted interfaces of bacterial
985 competition. *J Bacteriol* **198**: 2145–2155.
986 Taylor, N.M.I., van Raaij, M.J., and Leiman, P.G. (2018) Contractile injection systems of
987 bacteriophages and related systems. *Mol Microbiol* **108**: 6–15.

bioRxiv preprint doi: https://doi.org/10.1101/2020.07.24.220178; this version posted July 24, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

988 Taylor, R.K., Miller, V.L., Furlong, D.B., and Mekalanos, J.J. (1987) Use of *phoA* gene
989 fusions to identify a pilus colonization factor coordinately regulated with cholera toxin.
990 *Proc Natl Acad Sci USA* **84**: 2833–2837.
991 Unterweger, D., Kitaoka, M., Miyata, S.T., Bachmann, V., Brooks, T.M., Moloney, J., *et al*.
992 (2012) Constitutive type VI secretion system expression gives *Vibrio cholerae* intra- and
993 interspecific competitive advantages. *PLoS One* **7:** e48320.
994 Unterweger, D., Kostiuk, B., Otjengerdes, R., Wilton, A., Diaz-Satizabal, L., and Pukatzki, S.
995 (2015) Chimeric adaptor proteins translocate diverse type VI secretion system effectors
996 in *Vibrio cholerae*. *EMBO J* **34**: 2198–2210.
997 Unterweger, D., Miyata, S.T., Bachmann, V., Brooks, T.M., Mullins, T., Kostiuk, B., *et al*.
998 (2014) The *Vibrio cholerae* type VI secretion system employs diverse effector modules
999 for intraspecific competition. *Nat Commun* **5**: 3549.
1000 Veening, J.W. and Blokesch, M. (2017) Interbacterial predation as a strategy for DNA
1001 acquisition in naturally competent bacteria. *Nat Rev Microbiol* **15**: 621–629.
1002 Waldor, M.K. and Mekalanos, J.J. (1996) Lysogenic conversion by a filamentous phage
1003 encoding cholera toxin. *Science* **272**: 1910–1914.
1004 Wattam, A.R., Davis, J.J., Assaf, R., Boisvert, S., Brettin, T., Bun, C., *et al*. (2017)
1005 Improvements to PATRIC, the all-bacterial bioinformatics database and analysis
1006 resource center. *Nucleic Acids Res* **45**: D535–D542.
1007 Watve, S.S., Thomas, J., and Hammer, B.K. (2015) CytR is a global positive regulator of
1008 competence, type VI secretion, and chitinases in *Vibrio cholerae*. *PLoS One* **10:**
1009 e0138834**.**
1010 Weill, F.-X., Domman, D., Njamkepo, E., Tarr, C., Rauzier, J., Fawal, N., *et al*. (2017)
1011 Genomic history of the seventh pandemic of cholera in Africa. *Science* **358**: 785–789.
1012 Weill, F.X., Domman, D., Njamkepo, E., Almesbahi, A.A., Naji, M., Nasher, S.S., *et al*.

bioRxiv preprint doi: https://doi.org/10.1101/2020.07.24.220178; this version posted July 24, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

(2019) Genomic insights into the 2016–2017 cholera epidemic in Yemen. *Nature* **565**: 230–233.

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

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

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

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

Zoued, A., Brunet, Y.R., Durand, E., Aschtgen, M.S., Logger, L., Douzi, B., *et al.* (2014) Architecture and assembly of the Type VI secretion system. *Biochim Biophys Acta - Mol Cell Res* **1843**: 1664–1673.

bioRxiv preprint doi: https://doi.org/10.1101/2020.07.24.220178; this version posted July 24, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

Figure legends

Figure 1: Defense against amoebal predation by a subset of *V. cholerae* strains. (A, D) Amoebal predation was scored using *D. discoideum* grazing assays in which formed plaques on bacterial lawns were enumerated. Plaque numbers are indicated relative to those formed on a lawn of *K. pneumoniae*, which served as a positive control. Bar plots represent the average of at least three independent biological replicates (±SD). Statistical significance is indicated (n.s., not significant; **** $p < 0.0001$). (B and C) Hemolytic activity was tested on blood agar plates. Pandemic *V. cholerae* strain A1552 as well as all environmental isolates (B) or a representative subset together with their respective *hlyA*-minus derivatives (C) were assessed for hemolysis.

Figure 2: Constitutive T6SS activity linked to an ACD-containing effector inhibits amoebal grazing. (A, C, F) Bacterial killing assays using *E. coli* as prey. Numbers of surviving prey are depicted on the Y-axis (CFU/ml). Statistical significance in panel A is shown above each strains' bar and calculated relative to the T6SS-silent negative control strain A1552. #, for these strains, the killing activity was only reduced in one of the three independent experiments. (B) T6SS activity in representative environmental strains. Hcp detection in WT and Δ*vipA* mutants of representative environmental isolates. Intracellular (pellet) and secreted (supernatant) Hcp were assessed by immunoblotting using Hcp-directed antibodies. Detection of σ70 served as a loading control. (D, G) T6SS- and ACD-dependency of the anti-amoebal defense. Plaque formation by *D. discoideum* on bacterial lawns formed by representative *V. cholerae* WT, *vipA* derivatives (D and G) and ACD-minus (G) strains. Details as in Fig. 1. The toxigenic non-pandemic strain ATCC25872 and its site-directed mutant served as control in panel G. (E) Simplified scheme of the T6SS. The actin crosslinking domain (ACD) consists of a C-terminal extension of the VgrG1 tip protein and

bioRxiv preprint doi: https://doi.org/10.1101/2020.07.24.220178; this version posted July 24, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

this multidomain protein is encoded by the *vgrG1* locus (shown on the right). Removal of the ACD-encoding sequence was accomplished through site-directed integration of a stop codon concomitantly with an *aph* selective marker. Bar plots in all panels represent the average of at least three independent biological replicates (±SD). < dl, below detection limit. Statistical significance is indicated (n.s., not significant; * $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$; **** $p < 0.0001$; for panel A, each sample was compared to the A1552 control).

Figure 3: T6SS effector/immunity typing scheme of pandemic and environmental *V. cholerae*. The previously published phylogenetic tree was freely adapted from Keymer *et al*. 2007 (not fully to scale). The clade color code is depicted in the background. The E/I type of each T6SS cluster (large cluster as well as auxiliary (aux) clusters 1, 2, 3 and 6) are schematized for each strain. Large black arrows symbolize *vgrG* genes, which were classified as structural (black) or evolved (colored tips, to represent the different types of C-terminal effector domains). Large and small colored arrows represent effector and immunity genes, respectively, according to the color code indicated in the legend below the scheme. Immunity genes that are not adjacent to a putative effector gene are considered as orphan immunity loci. * depicts an orphan immunity gene from strain L6G whose gene product was slightly below the amino acid identity threshold of 30% (20.3%) relative to C-type immunity proteins. # depicts an orphan immunity gene that is a homolog to *bona fide* immunity genes in other genomes. Effector genes in auxiliary clusters 1 and 2 of strain W10G (marked in figure with a pattern) were wrongly annotated in the PacBio genome sequence due to a frameshift sequencing artifact splitting each gene into two ORFs; this sequencing error was corrected by Sanger sequencing and the corrected single gene is shown in this figure.

bioRxiv preprint doi: https://doi.org/10.1101/2020.07.24.220178; this version posted July 24, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

Figure 4: Competitiveness among environmental strains. Graphs in all panels represent bacterial killing assays as described in Fig. 2 using the predator strains as indicated on the X-axes and as prey: (A) the pandemic *V. cholerae* strain A1552; (B) representative *vipA*-minus (ΔT6SS) mutants of the environmental isolates as shown above the graph; or (C to Q) each environmental isolate as shown in the graph title for each panel and on the Y-axis. Plots represent the average of three independent biological replicates (±SD). < dl, below the detection limit. Statistical significance is indicated (n.s., not significant; * $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$; **** $p < 0.0001$).

bioRxiv preprint doi: https://doi.org/10.1101/2020.07.24.220178; this version posted July 24, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

Supporting tables and figure legends

Table S1. *Vibrio cholerae*, *Escherichia coli* and *Klebsiella pneumoniae* strains used in this study.

Table S2. Information of the long-read whole genome sequencing data and assemblies of NCBI BioProject PRJNA633476.

Table S3. BlastP-predicted T6SS effector and immunity proteins[1] of the environmental *V. cholerae* isolates.

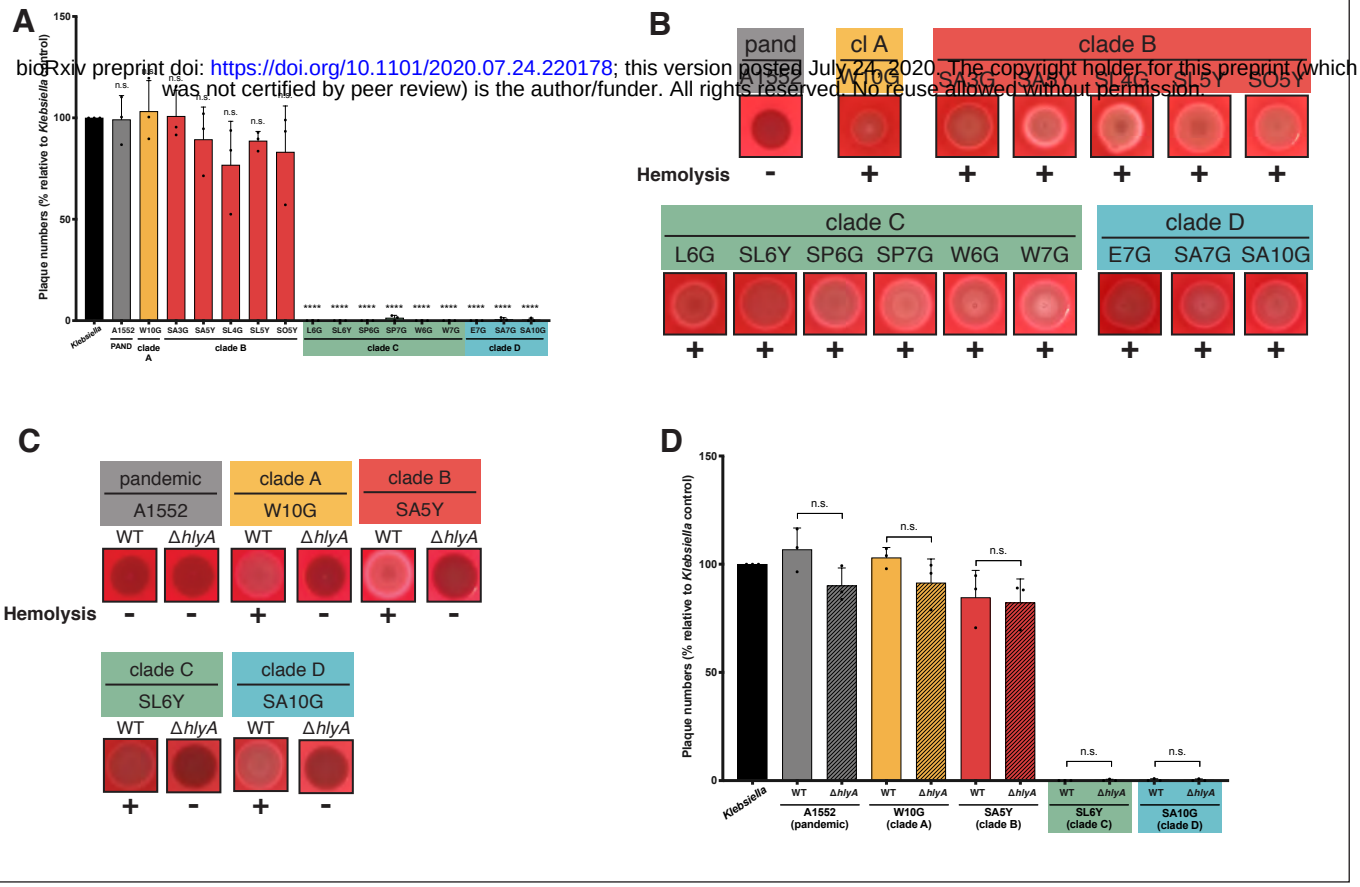
Table S4. Matrices of % identity (percentage of residues that are identical) among the effector and immunity proteins harbored in the T6SS large cluster from the *V. cholerae* environmental strains and the pandemic A1552 strain.

Table S5. Matrices of % identity (percentage of residues that are identical) among the effector and immunity proteins harbored in the T6SS auxiliary cluster 1 from the *V. cholerae* environmental strains and the pandemic A1552 strain.

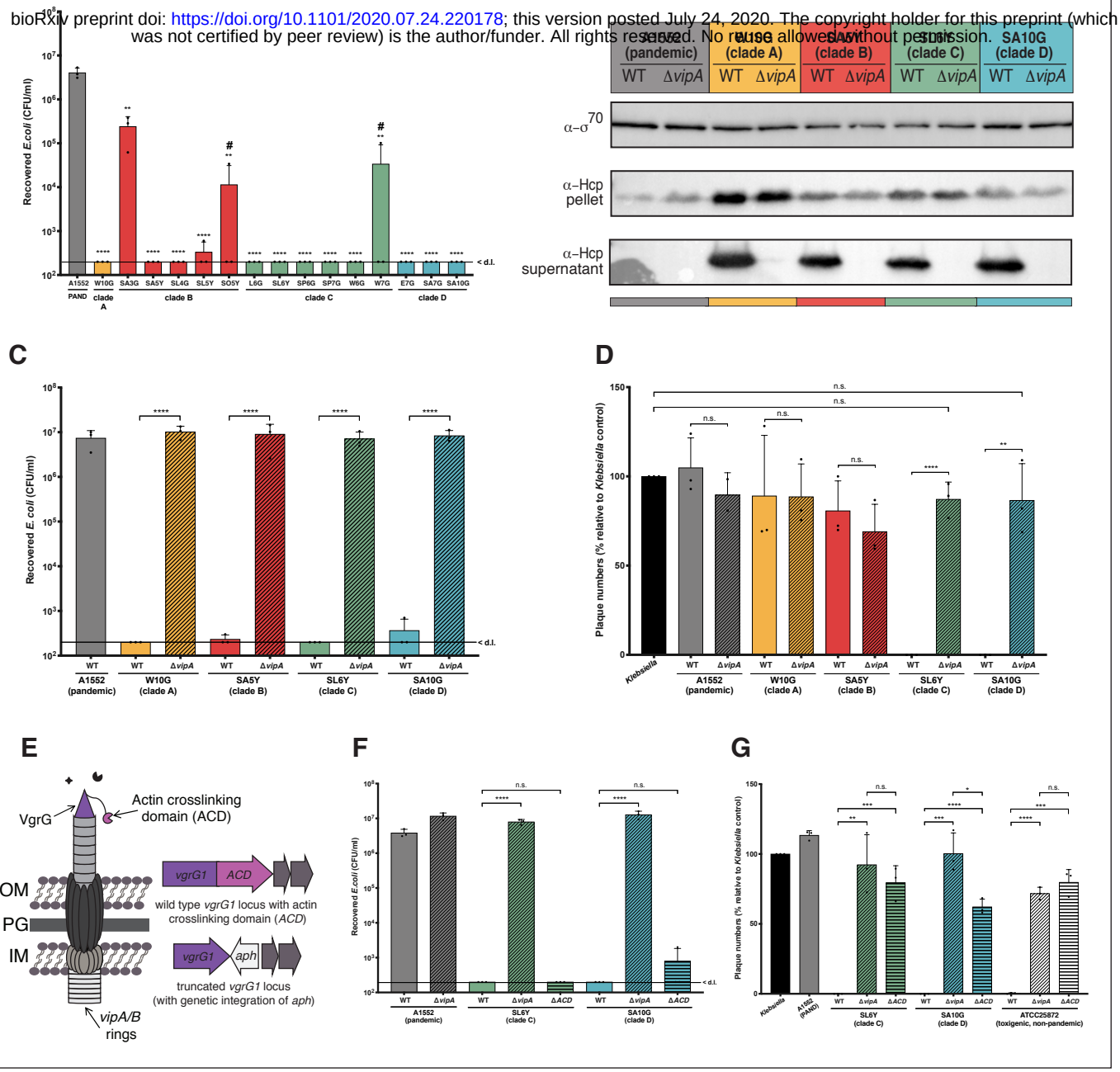
Table S6. Matrices of % identity (percentage of residues that are identical) among the effector and immunity proteins harbored in the T6SS auxiliary cluster 2 from the *V. cholerae* environmental strains and the pandemic A1552 strain.

Supplementary Fig. S1: Scheme of T6SS auxiliary cluster 6 of strain SP7G. The 4.1-kb long auxiliary T6SS cluster carried on the chromosome 2 of strain SP7G (locus tags VC-SP7G_03229 to VC-SP7G_03234 from CP053809) is depicted. Gene symbols are explained below the scheme.

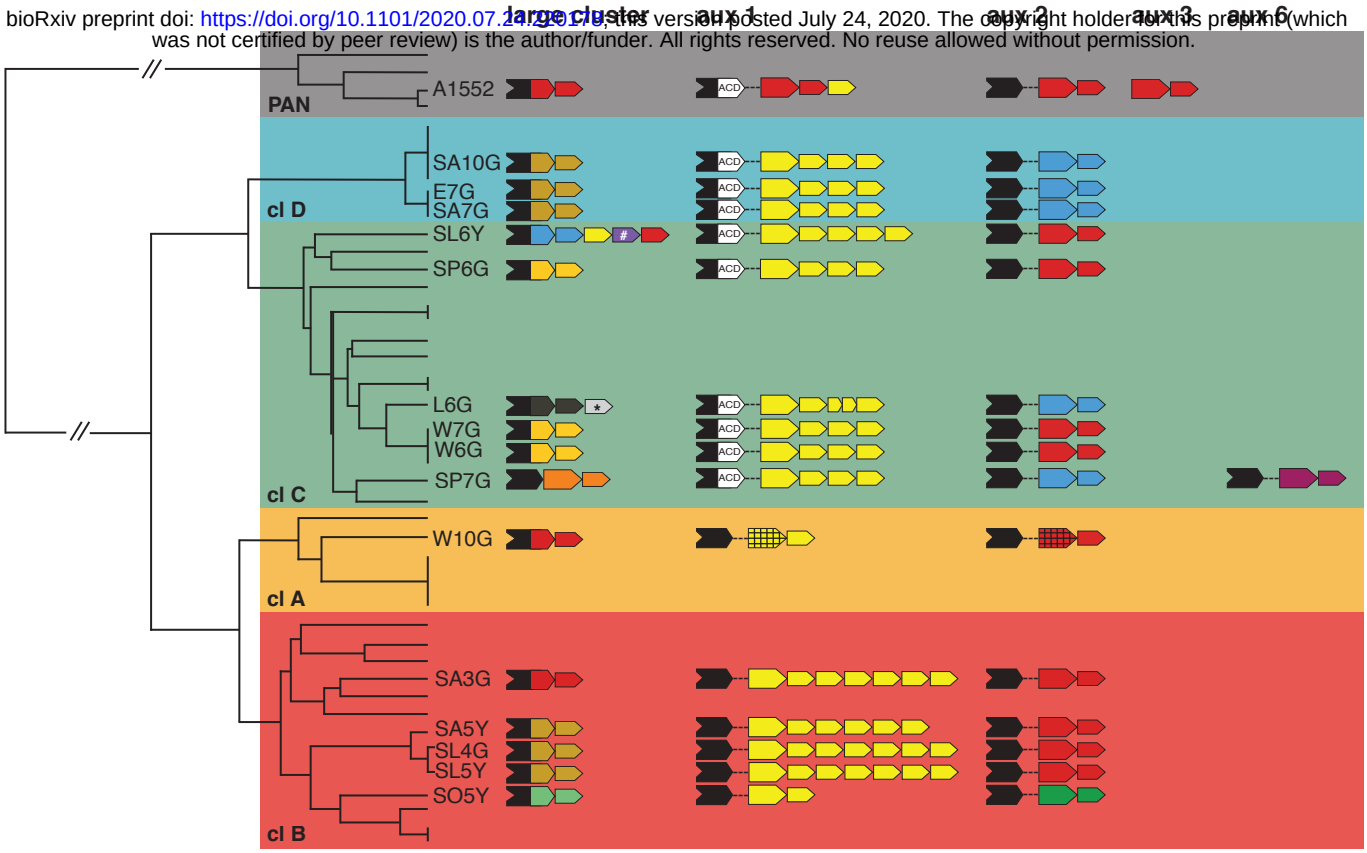
bioRxiv preprint doi: https://doi.org/10.1101/2020.07.24.220178; this version posted July 24, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.



bioRxiv preprint doi: https://doi.org/10.1101/2020.07.24.220178; this version posted July 24, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.



bioRxiv preprint doi: https://doi.org/10.1101/2020.07.24.220178; this version posted July 24, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.



Effector and immunity modules in T6SS clusters

structural VgrG
evolved VgrG
effector
immunity/ orphan
ORF corrected by Sanger sequencing
---- additional genes

large cluster
A E I
B G J
C H K

aux cluster 1
VgrG1
VgrG1 ACD
A
C

aux cluster 2
A
D
E

aux cluster 3
A

aux cluster 6

bioRxiv preprint doi: https://doi.org/10.1101/2020.07.24.220178; this version posted July 24, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.