Exploring Mobile Genetic Elements in *Vibrio cholerae*

Natália C. Drebes Dörr [1,2], Alexandre Lemopoulos [1], Melanie Blokesch [1,*]

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

[2]Present address: Departamento de Genética, Evolução, Microbiologia e Imunologia, Instituto de Biologia, Universidade Estadual de Campinas (Unicamp), Campinas, SP, Brazil

*Corresponding author: E-mail: melanie.blokesch@epfl.ch.

Accepted: April 25, 2025

Abstract

Members of the bacterial species *Vibrio cholerae* are known both as prominent constituents of marine environments and as the causative agents of cholera, a severe diarrheal disease. While strains responsible for cholera have been extensively studied over the past century, less is known about their environmental counterparts, despite their contributions to the species' pangenome. This study analyzed the genome compositions of 46 *V. cholerae* strains, including pandemic and nonpandemic, toxigenic, and environmental variants, to investigate the diversity of mobile genetic elements (MGEs), embedded bacterial defense systems, and phage-associated signatures. Our findings include both conserved and novel MGEs across strains, pointing to shared evolutionary pathways and ecological niches. The defensome analysis revealed a wide array of anti-phage/antiplasmid mechanisms, extending well beyond the traditional CRISPR-Cas and restriction-modification systems. This underscores the dynamic arms race between *V. cholerae* and MGEs and suggests that nonpandemic strains may act as reservoirs for emerging defense strategies. Moreover, the study showed that MGEs are integrated into genomic hotspots, which may serve as critical platforms for the exchange of defense systems, thereby enhancing *V. cholerae's* adaptive capabilities against phage attacks and other invading MGEs. Overall, this research offers new insights into *V. cholerae's* genetic complexity and potential adaptive strategies, offering a better understanding of the differences between environmental strains and their pandemic counterparts, as well as the possible evolutionary pathways that led to the emergence of pandemic strains.

Key words: *Vibrio cholerae*, horizontal gene transfer, mobile genetic elements, defense systems, phages, bacterial immunity.

Significance

Vibrio cholerae, the causative agent of cholera, has been extensively studied in its pandemic form, but nonpandemic strains remain understudied, creating gaps in understanding their genetic diversity and defense mechanisms. This study analyzed the genomes of 46 *V. cholerae* strains, revealing a broad range of mobile genetic elements and known or predicted defense systems. These findings suggest that nonpandemic strains could be a crucial reservoir for emerging defense strategies, offering new insights into the adaptive capabilities of *V. cholerae* and potentially informing our understanding of the evolutionary paths that lead to pandemic strains.

Introduction

Classifying prokaryotic genomes into specific units, such as species, is often a complex task due to their inherent genomic fluidity (Frost et al. 2005; Doolittle and Papke 2006; Toussaint and Chandler 2012; Shapiro and Polz 2014). Comparative studies have shown that genomic variation within the same prokaryotic species is less about sequence variability and more about differences in gene content (Doolittle and Papke 2006; Innamorati et al. 2020). These variations in gene content are largely driven by mobile genetic elements (MGEs), including genomic islands, plasmids, phages, transposons, insertion sequence (IS) elements, integrative and conjugative elements (ICEs), and integrons, all of which play a crucial role in bacterial evolution (Arnold et al. 2022). Indeed, the mobilome, representing the entire set of MGEs within a genome, is a significant component of the prokaryotic genome landscape. Although some MGEs can change their chromosomal position within the same cell, they are most commonly transferred between different organisms' genomes via horizontal gene transfer (HGT). This transfer typically occurs through mechanisms like transformation, conjugation, or transduction (Ochman et al. 2000; Frost et al. 2005; Toussaint and Chandler 2012; Arnold et al. 2022).

The general term "genomic island" (GI) refers to MGEs that have become established within a genome (Dobrindt et al. 2004). These regions typically exhibit distinct characteristics compared with the surrounding genome, such as an unusual GC content, flanking repeat sequences, and preferential integration near tRNA genes or other highly conserved genomic regions. Despite their chromosomal integration, many GIs retain the ability to mobilize due to the presence of integrase-, transposase-, or conjugation-related genes (Dobrindt et al. 2004; Juhas et al. 2009; Arnold et al. 2022). Beyond their mobility, GIs often encode genetic features that confer adaptive advantages to their bacterial hosts. For instance, some GIs carry virulence-related genes that facilitate the transition of bacteria to a pathogenic lifestyle (Frost et al. 2005; Doolittle and Papke 2006; Ahmed et al. 2008). Others contribute to a range of adaptive traits, including the degradation of specific compounds, niche adaptation, symbiosis, interbacterial competition, and defense mechanisms against foreign DNA and phages (Dobrindt et al. 2004; Frost et al. 2005; Juhas et al. 2009).

Vibrio cholerae serves as an exemplary model to study the evolutionary impact of MGEs. This gram-negative bacterium is the causative agent of cholera, a severe diarrheal disease. However, most *V. cholerae* strains, and especially environmental isolates, are not implicated in the disease cholera. Notably, out of the more than 200 *V. cholerae* serogroups, only two (O1 and O139) are associated with historical and ongoing cholera pandemics (Clemens et al. 1991; Faruque et al. 1998; Cottingham et al. 2003).

Horizontally acquired MGEs play a crucial role in enabling *V. cholerae* to become toxigenic, allowing it to cause cholera disease. Furthermore, they have also contributed to the classification of lineages from past and current cholera pandemics and allow us to distinguish distinct waves within the ongoing seventh pandemic (Dziejman et al. 2002; Mutreja et al. 2011; Robins and Mekalanos 2014; Boyd et al. 2015). Briefly, cholera-causing *V. cholerae* strains are characterized by the presence of *Vibrio* pathogenicity island 1 (VPI-1, also known as TCP island), which encodes the toxin-coregulated pilus (TCP). TCP is vital for intestinal colonization (Taylor et al. 1987; Karaolis et al. 1998) and serves as the receptor for the CTX bacteriophage (CTXΦ) (Waldor and Mekalanos 1996). The CTXΦ prophage harbors genes for the production of cholera toxin (CTX), the primary virulence factor responsible for the hallmark symptom of acute diarrhea (Waldor and Mekalanos 1996). In addition to these primary and virulence-encoding MGEs, cholera-causing strains also contain VPI-2. This island includes the *nan-nag* region, which plays a role in sialic acid metabolism (Jermyn and Boyd 2005). Additionally, VPI-2 contains a cluster of phage-related genes and operons that encode established (DdmDE, TgvAB, T1RM) or predicted phage and DNA defense systems (Jermyn and Boyd 2005; Murphy and Boyd 2008; Boyd et al. 2015; Jaskólska et al. 2022; Gomez and Waters 2024; Vizzarro et al. 2024). Moreover, pandemic strains responsible for the current 7th cholera pandemic, often referred to as 7th pandemic El Tor strains or 7PET, are characterized by two specific MGEs, known as *Vibrio* seventh pandemic islands I and II (VSP-I, VSP-II) (Dziejman et al. 2002; O'Shea et al. 2004; Davies et al. 2012; Boyd et al. 2015). These genetic markers are pivotal for differentiating these strains from those responsible for previous pandemics.

Although the mobilome of pandemic *V. cholerae* strains has been extensively characterized, the genetic diversity within nonpandemic isolates, whether pathogenic or environmental, is still not well-understood. Previous studies primarily used PCR-based methods, focusing on confirmatory- rather than discovery-driven analyses (Rivera et al. 2001; Faruque et al. 2003, 2004; Gennari et al. 2012; Bernardy et al. 2016). In contrast, only a few studies have utilized whole-genome sequences (WGS) for a more comprehensive investigation of individual GIs in environmental *V. cholerae* isolates (Chun et al. 2009; Rapa et al. 2015; Labbate et al. 2016).

In this study, we explored the mobilome diversity of nonpandemic strains and compared it to that of the 7th pandemic clade of *V. cholerae*. We analyzed high-quality genome sequencing data using various methods to identify GIs within these strains. Additionally, we employed tools to detect defense systems, which are often found in MGEs (Benler et al. 2021; LeGault et al. 2021; Rousset et al. 2022; Vassallo et al. 2022; Hochhauser et al. 2023).

Our findings revealed significant variability in the horizontal gene pool of these strains, suggesting a dynamic genetic composition. Our analysis also showed that gene operons were present in regions of genomic plasticity (RGP), often in varied combinations. Notably, some of these operons are also present in GIs associated with pandemic strains, suggesting a potential reservoir for novel features within nonpandemic strains of *V. cholerae*.

Results and Discussion

Strain Selection and Phylogenetic Analysis

For our analysis, we compiled a total of 46 WGS, consisting of 45 *V. cholerae* strains and one *V. mimicus* strain, which was, as expected, placed as an outgroup of the phylogenetic tree, as previously described (Karaolis et al. 1995), using the minimal ancestral deviation (MAD) method (Tria et al. 2017) (supplementary tables S1 and S2, Supplementary Material online). This collection encompassed genomes from eight 7PET strains (supplementary table S1, Supplementary Material online), which were used as controls. These 7PET strains were isolated at various stages of the ongoing 7th cholera pandemic: early strains from the 1970s in Bangladesh and Bahrain (N16961, P27459, and E7946) (Miller et al. 1989; Pearson et al. 1993; Heidelberg et al. 2000; Matthey et al. 2018; Stutzmann and Blokesch 2020); two strains from the 1990s Peruvian cholera outbreak (A1552 and C6706) (Yildiz and Schoolnik 1998; Stutzmann and Blokesch 2016, 2020; Matthey et al. 2018); and three recent strains from 2009 to 2012 isolated in the Democratic Republic of the Congo (DRC193A, DRC052, and DRC072) (Lemopoulos et al. 2024). Additionally, the dataset included a prepandemic strain, MAK757, isolated in 1937 in Indonesia (Chattopadhyay et al. 1993), considered to be part of the so-called El Tor progenitor group (Mutreja et al. 2011; Boucher 2016). The assembled WGS of this strain was deposited as part of this work (see Materials and Methods for details).

Our phylogenetic analysis distinguished the 7PET clade as a strongly supported monophyletic group (Fig. 1a). As expected, the prepandemic MAK757 strain was positioned as progenitor of the El Tor group (Mutreja et al. 2011; Boucher 2016). In contrast to the distinct lineage of the pandemic clade, nonpandemic strains did not form a monophyletic clade. Furthermore, the analysis did not reveal any specific branching pattern based on source of isolation, indicating that strains from environmental or clinical sources did not cluster distinctly. Also, no clear biogeographic signal was observed in the nonpandemic strains' groupings. However, a highly supported clade comprising strains originating from California (namely SA7G, E7G, SA10G, SP6G, DL4211, W7G, W6G, TP, SP7G, L6G, and SL6Y) was noted, although this did not form a monophyletic grouping with other Californian isolates such as SL5Y, SA5Y, and SIO. This phylogenetic resolution aligns with previous studies on the California strains that used comparative genome hybridization on microarrays (Keymer et al. 2007; Miller et al. 2007), confirming the well-supported clades A, B, and D, and revealing the polyphyletic nature of clade C.

Interestingly, strain RFB05, initially identified in NCBI as a *V. cholerae* strain (Genbank: GCA_008369625.1), demonstrated phylogenetic affinity toward the root and *V. mimicus*. A secondary phylogenetic analysis (genomes used are depicted in supplementary table S4, Supplementary Material online) suggested that strain RFB05 is likely misidentified and belongs to the *V. tarriae* species (Islam et al. 2022) rather than to *V. cholerae* (supplementary fig. S1a, Supplementary Material online).

Determination of Conserved Islands in Pandemic Strains as Proof of Concept

We next sought to determine the strains' GIs, as outlined in the methods section (supplementary table S3, Supplementary Material online). As a demonstration of our methodology's effectiveness, we were able to identify all well-known MGEs associated with pandemic *V. cholerae* strains, as shown in Fig. 1b. Notably, the prepandemic strain MAK757 lacked VSP-I and II, aligning with expectations (Fig. 1b). Sequence alignments and gene content analyses confirmed the expected high level of conservation of these GIs (i.e. VPI-1 and 2, VSP-I and II) across strains from various stages of the 7th cholera pandemic (supplementary figs. S1 and S2, Supplementary Material online). Interestingly, DRC072 and DRC052 strains carry an additional VSP-I copy on chromosome 2, near the *uhpC* gene (encoding an MFS transporter family glucose-6-phosphate receptor; supplementary fig. S1e, Supplementary Material online).

Our methods also detected MGEs specific to certain pandemic strains, like the WASA-1 prophage in South American strains (A1552 and C6706) integrated into the alanine aminopeptidase gene (*pepN*) (Mutreja et al. 2011; Garza et al. 2012; Adams et al. 2025) (Fig. 1b and supplementary fig. S2e, Supplementary Material online), and the STX-like ICE (ICE*Vch*Ind5; GQ463142) in DRC strains (Fig. 1b and supplementary fig. S2f, Supplementary Material online). SXT-like ICEs are common in recent clinical isolates from Asia and Africa (Wozniak et al. 2009) and integrate site-specifically into the *prfC* gene (Hochhut and Waldor 1999).

In our study, we identified several additional MGEs that showed a high degree of conservation across all pandemic *V. cholerae* strains (Fig. 1b). Notably, the Vc95 Retron, inserted next to gene *VC2384*, was almost identical in all examined pandemic strains (supplementary fig. S2g, Supplementary Material online). This retron plays a role in antiphage defense, comprising genes for a reverse

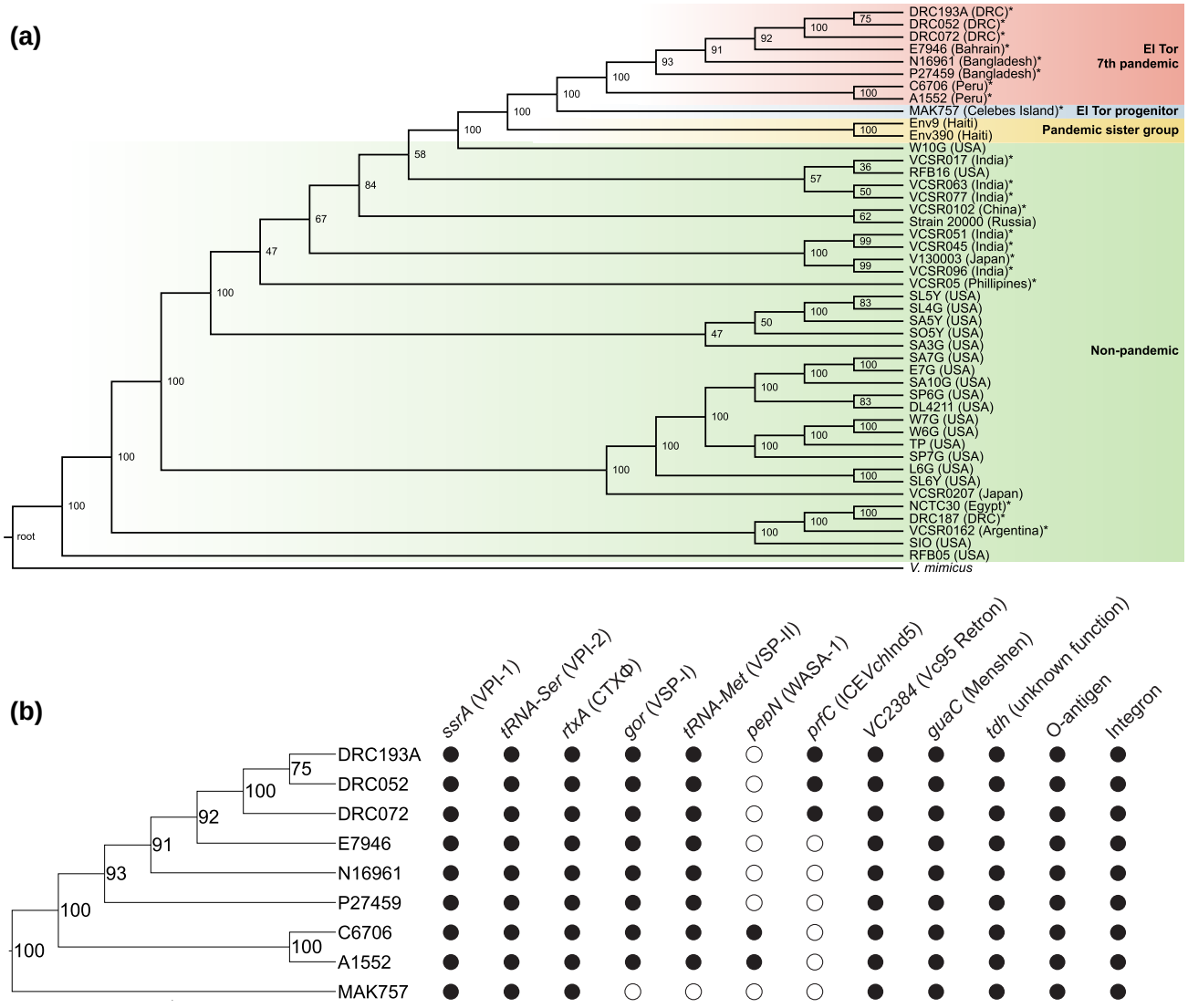

Fig. 1. Identification of MGEs in pandemic *V. cholerae* strains. a) This phylogenetic tree, based on 1,532 core genes, depicts the evolutionary relationship among 45 *V. cholerae* strains while *V. mimicus* (ATCC33655) was placed as an outgroup by the MAD. Statistical confidence was assessed with 100 bootstraps. 7PET strains (highlighted in red) and the prepandemic strain MAK757 (in blue), are clearly marked, whereas nonpandemic strains, including a pandemic sister group, are indicated in green and yellow, respectively. Strains isolated from patients are marked with an asterisk. b) A detailed view of the phylogenetic tree showing only the pandemic strains. Black circles indicate the presence, and white circles the absence, of known pandemic GIs in each strain. The insertion sites are indicated at the top, with the names of the MGEs displayed in brackets.

transcriptase, an ATPase, and an HNH endonuclease in addition to the noncoding RNA *msr-msd* (Millman et al. 2020; Bobonis et al. 2022). Moreover, every pandemic strain was found to harbor a compact island equipped with genes encoding a predicted Menshen antiphage defense system near the *guaC* gene (supplementary fig. S2h, Supplementary Material online), and another island populated with numerous hypothetical genes near the *tdh* gene (supplementary fig. S2i, Supplementary Material online), both of them located on chromosome 2. The specific insertion sites, *VC2384*, *guaC* and *tdh*, also contain MGEs in nonpandemic strains, as discussed below.

Mobilome Diversity in Nonpandemic *V. cholerae*

Focusing on nonpandemic *V. cholerae*, our analysis revealed the presence of MGEs across 32 insertion sites on chromosome 1 and 13 sites on chromosome 2 (Fig. 2a and b; supplementary tables S5 and S6, Supplementary Material online). Interestingly, all insertion sites housing MGEs in pandemic strains were also occupied in nonpandemic strains, with the exception of *pepN* (carrying the WASA-1 prophage in West African–South American pandemic strains) (Fig. 2a). This finding suggests that these genomic locations, and particularly tRNA/tmRNA sites as previously reported (Boyd et al. 2009), serve as

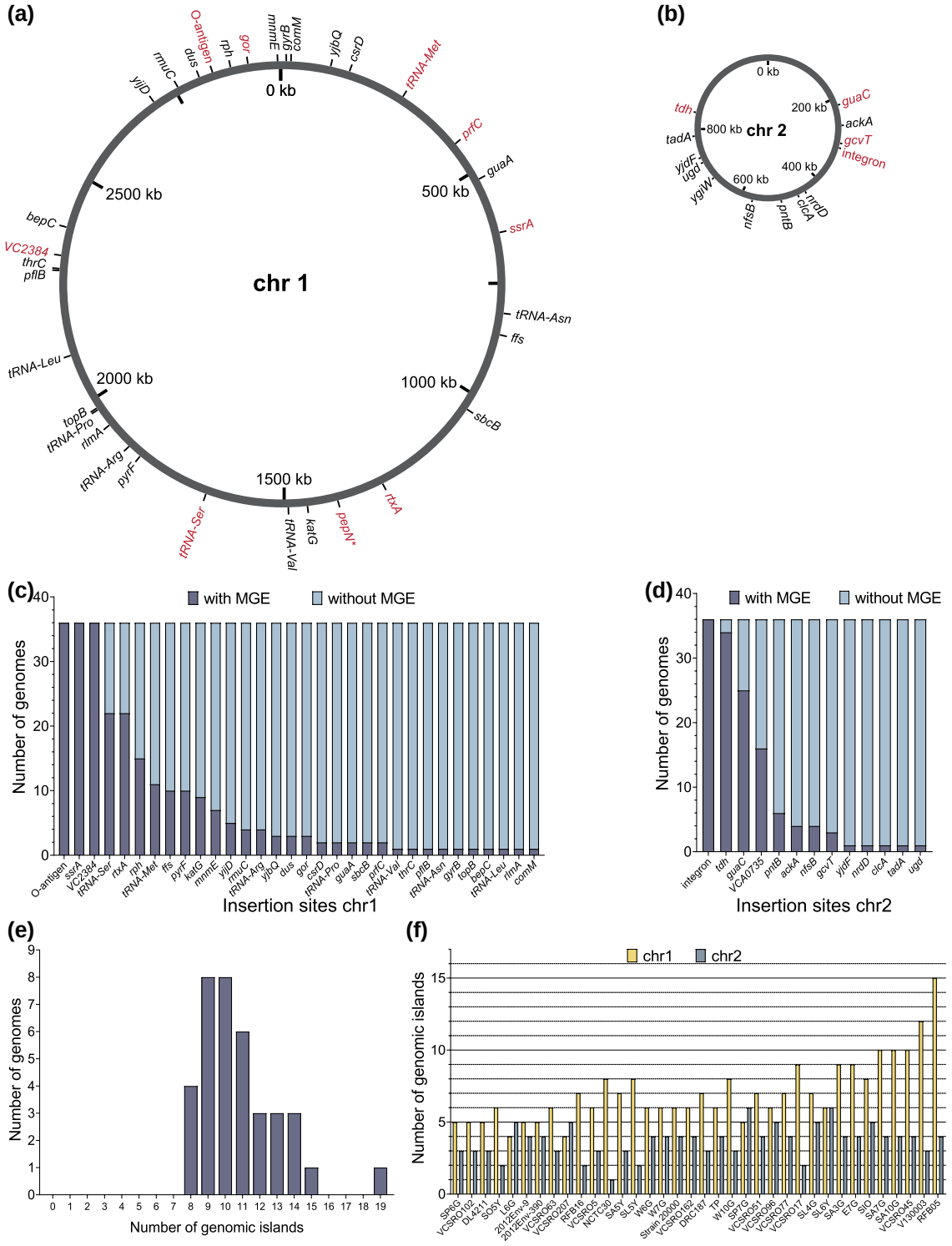

Fig. 2. RGP in nonpandemic *V. cholerae* strains. a, b) Scheme of *V. cholerae* chromosomes 1 (a) and 2 (b), displaying all insertion sites where MGEs are found in at least one strain. Insertion sites occupied in pandemic strains are marked in red, with those unique to pandemic strains indicated by an asterisk. The layout of these sites follows the genomic structure of pandemic strain A1552. For comprehensive details on all insertion sites, see supplementary tables S5 and S6, Supplementary Material online. c, d) Charts showing the presence of MGEs at insertion sites across chromosome 1 (c) and chromosome 2 (d) within the dataset of nonpandemic *V. cholerae* strains. For each site, the charts display the count of genomes with or without an MGE. e) The distribution graph showing the total count of GIs identified per genome in our nonpandemic strain dataset, highlighting *V. tarriae* RFB05 with 19 GIs as an outlier. f) Graph illustrating the quantity of MGEs located on chromosomes 1 and 2 for each nonpandemic *V. cholerae* strain examined.

preferred insertion sites, potentially due to conserved integration mechanisms.

The most frequently occupied insertion sites in nonpandemic strains included the genes *rjg*, *ssrA*, *VC2384,* and *tRNA-Ser* in chromosome 1, as well as the integron, *tdh*, and *guaC* in chromosome 2 (Fig. 2c and d; supplementary table S7, Supplementary Material online). Specifically, while our study did not focus on these genetic elements, our analysis detected the O-antigen cluster located near the *rjg* gene on chromosome 1 (Murase et al. 2022) and a sizable integron on chromosome 2 (Fig. 2c and d). Additionally, the tmRNA gene *ssrA*, a common target for MGE integration due to independent evolution of many integrases toward this site (Williams 2003; Boyd et al. 2009), was found to be occupied by MGEs in all examined nonpandemic strains. Similarly, tRNA genes, favored by many MGEs for their conserved sequences (Williams 2002; Boyd et al. 2009), saw the *tRNA-Ser* site occupied by MGEs in over half of our nonpandemic strain dataset (Fig. 2c).

MGE abundance varied from 8 to 15 GIs per genome, with most strains containing 10 GIs (taking into account the specific cutoffs we set out to consider a *bona fide* GI) (Fig. 2e). Most of the islands were found on the larger chromosome 1 (3 Mbp compared with 1 Mbp of chromosome 2), but some strains (L6G, VCSRO207, SP7G) had more GIs on chromosome 2 (Fig. 2f). This distribution pattern underscores the complex nature of genomic plasticity within *V. cholerae*, revealing the potential for diverse genetic configurations across different strains.

Secretion Systems Encoded in *V. cholerae*'s Mobilome

Our findings highlight a remarkable variability in gene content among MGEs located across a wide range of insertion sites within *V. cholerae* genomes. These MGEs harbor gene clusters involved in a spectrum of biological processes, from host-pathogen interactions and defense mechanisms to strategies for interbacterial competition, metabolism, and the production of cellular appendages (supplementary figs. S3 to S12, Supplementary Material online). Such diversity highlights the crucial role these gene clusters likely play in the environmental adaptation of *V. cholerae*, enabling the bacterium to navigate the diverse selective pressures encountered in its habitats. Interestingly, the presence of these gene clusters in nonpandemic strains, some of which have been isolated from human patients, suggests their potential significance in mediating interactions with the human host. This observation underscores the complex ecology of *V. cholerae*, which bridges both environmental reservoirs and human populations.

A notable finding was the widespread presence of secretion systems encoded within these GIs. Secretion systems, ranging from Type I to Type XI (T1SS to T11SS), are frequently found in elements that can be horizontally transferred between organisms (Dobrindt et al. 2004). These systems are composed of multiprotein complexes that transport a variety of substrates, including proteins and DNA, out of the cell. Such substrates are crucial for bacterial interactions with their surroundings, playing key roles in environmental adaptation and response (Costa et al. 2015). Our data revealed a significant number of strains equipped with auxiliary T6SS and/or T3SS gene clusters.

The T6SS functions as a sophisticated antibacterial and antieukaryotic arsenal, with its architecture in *V. cholerae* encoded by at least three gene clusters (Pukatzki et al. 2006; Boyer et al. 2009). The primary large cluster codes for the structural framework of the system, whereas the two auxiliary clusters, Aux1 and Aux2, equip the system with Hcp and VgrG proteins that form the spear-like apparatus, complemented by adaptor proteins and effector/immunity (E/I) pairs for precise target engagement and protection against self-intoxication (Crisan and Hammer 2020). Pandemic *V. cholerae* strains possess an additional auxiliary cluster, Aux3, enhancing their competitive edge with an extra set of E/I genes (Altindis et al. 2015; Hersch et al. 2020; Santoriello et al. 2020). Our analysis effectively pinpointed the presence of the Aux3 cluster across all examined pandemic strains, located at the *gcvT* insertion site, notable for its relatively small size under 10 kb. Their shorter size would exclude these islands from our analysis, but we kept them nonetheless for comparative reasons due to the identification of the larger (37-kb in length) version of this island, exhibiting mobile and prophage-like characteristics in nonpandemic strains 2012Env-9, 2012Env-390, and Strain 20000 (Fig. 3 and supplementary fig. S3, Supplementary Material online). This finding aligns with a previous report on Aux3 (Santoriello et al. 2020). Moreover, a significant portion of our dataset included strains containing the T6SS Aux4 and 5 clusters (Labbate et al. 2016; Crisan et al. 2019), reinforcing previous work (Otto et al. 2024). Aux4 clusters were found in islands inserted at *ssrA* (supplementary fig. S4, Supplementary Material online) and *tRNA-Ser* (Fig. 4a). Aux5 clusters, on the other hand, were found in MGEs inserted near the *tdh* or *pntB* genes on chromosome 2 (supplementary figs. S5 and S6, Supplementary Material online).

While the T6SS is primarily used for interbacterial warfare, the T3SS serves as a conduit between the bacterial cell and its host, enabling the transfer of effector proteins that can modulate host cellular responses (for review, see [Büttner 2012; Galán and Waksman 2018 ; Wagner et al. 2018]). We found that roughly one-third (13 out of the total number) of strains in our collection of nonpandemic *V. cholerae* harbored a T3SS within a GI located at the *tRNA-Ser* locus, a site typically associated with VPI-2 in pandemic strains. Interestingly, 11 of these 13 strains also incorporated the *nan-nag* region, a key component of VPI-2, within the same GI (Fig. 4a), confirming previous work (Murphy and Boyd 2008; Carpenter et al. 2017).

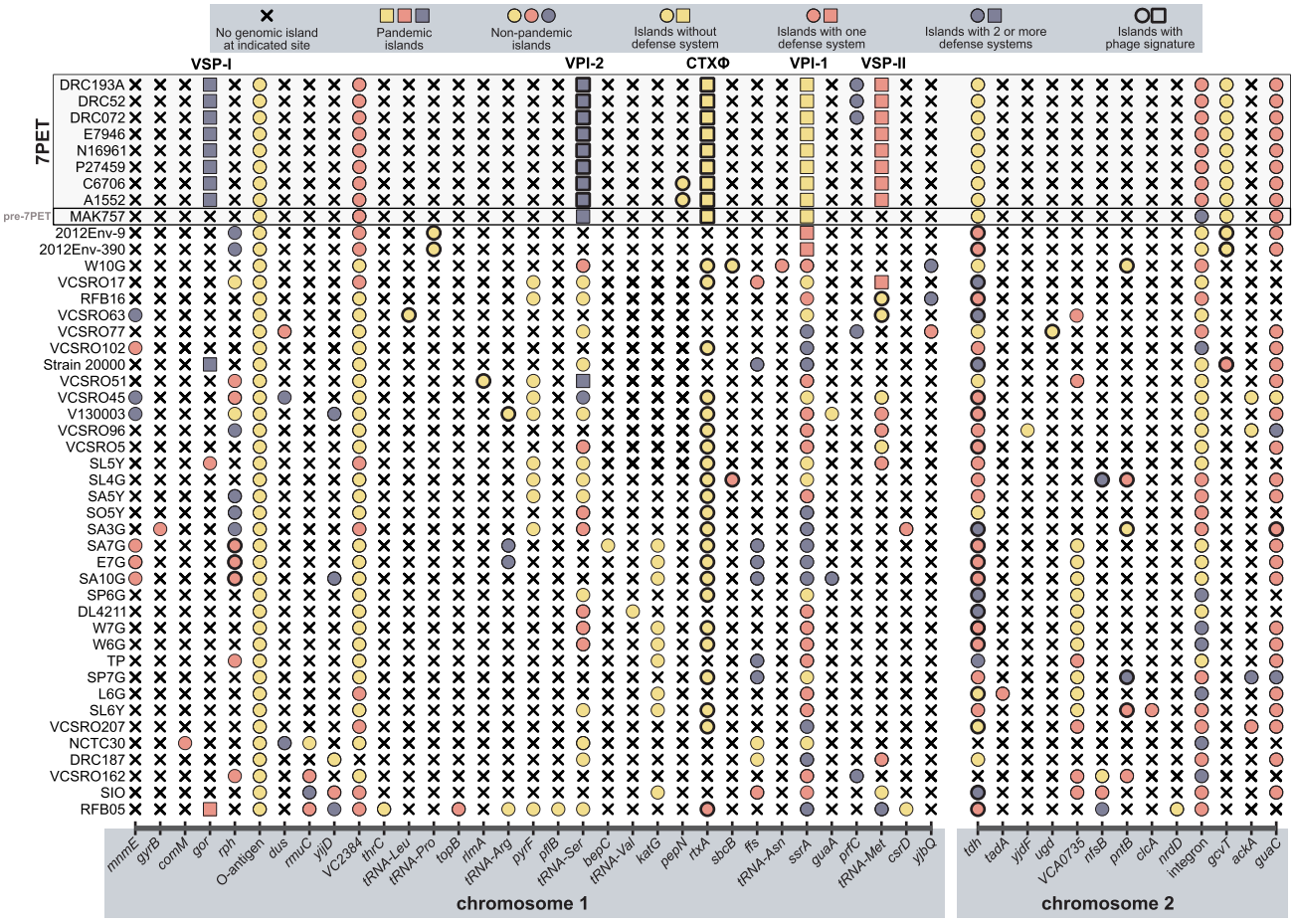

Fig. 3. Distribution of GIs and defense mechanisms in *V. cholerae*. This figure maps out the locations of MGEs across *V. cholerae* strains, both pandemic and nonpandemic. The *x* axis indicates the specific insertion sites where MGEs have been identified. Strain alignment on the left follows the order presented in the phylogenetic tree from Fig. 1. Each circle represents a GI, with pandemic-specific islands marked by squares. Color-coding identifies islands without antiphage/antiplasmid defense systems (yellow) and those with one or more defense systems (in red and gray). MGEs with phage-related signatures are highlighted with a bold outline.

These findings highlight the significance of T3SS islands in the genetic landscape of *V. cholerae*, particularly given their notable prevalence in our dataset, where over half of the strains with an island at the *tRNA-Ser* site possess a T3SS. Indeed, the T3SS, which shares a high degree of similarity with the T3SS2α of *V. parahaemolyticus*, has been extensively studied in the nonpandemic clinical CTX-minus isolate AM-19226 (Dziejman et al. 2005). This system translocates at least 13 effector proteins that interfere with various host pathways, inducing effects such as actin rearrangement and contributing to severe diarrhea in infant rabbits (Miller et al. 2019). The distribution of these islands appears to be widespread, with strains isolated from various global locations over different time periods and from a range of sources (supplementary table S2, Supplementary Material online). Notably, one such strain, NCTC 30, was isolated from a patient in Egypt in 1916 (Dorman et al. 2019), making it the oldest strain in our dataset and illustrating the long-standing presence of T3SS within the *V. cholerae* species.

Circulation of Pandemic-Like Islands in Nonpandemic Strains

Our analysis revealed the presence of nearly identical VPI-1 islands in nonpandemic strains 2012Env-9 and 2012Env-390, isolated from Haiti in 2012 (Fig. 3 and supplementary fig. S4; supplementary table S2, Supplementary Material online) (Azarian et al. 2016, 2014). The islands in these nontoxigenic O1 environmental isolates, previously considered as part of a pandemic sister group (Fig. 1a) (Boucher 2016), showed remarkable similarity to the *bona fide* VPI-1 found in pandemic strains but additionally carried genes encoding a Dnd antiphage defense system alongside the TCP and Acf clusters (supplementary fig. S4, Supplementary Material online). Interestingly, the Dnd gene cluster, but not the TCP/Acf clusters, was frequently identified within GIs located at the *ssrA* site in other environmental *V. cholerae* strains (L6G, TP, W10G, W7G, W6G; supplementary fig. S4, Supplementary Material online).

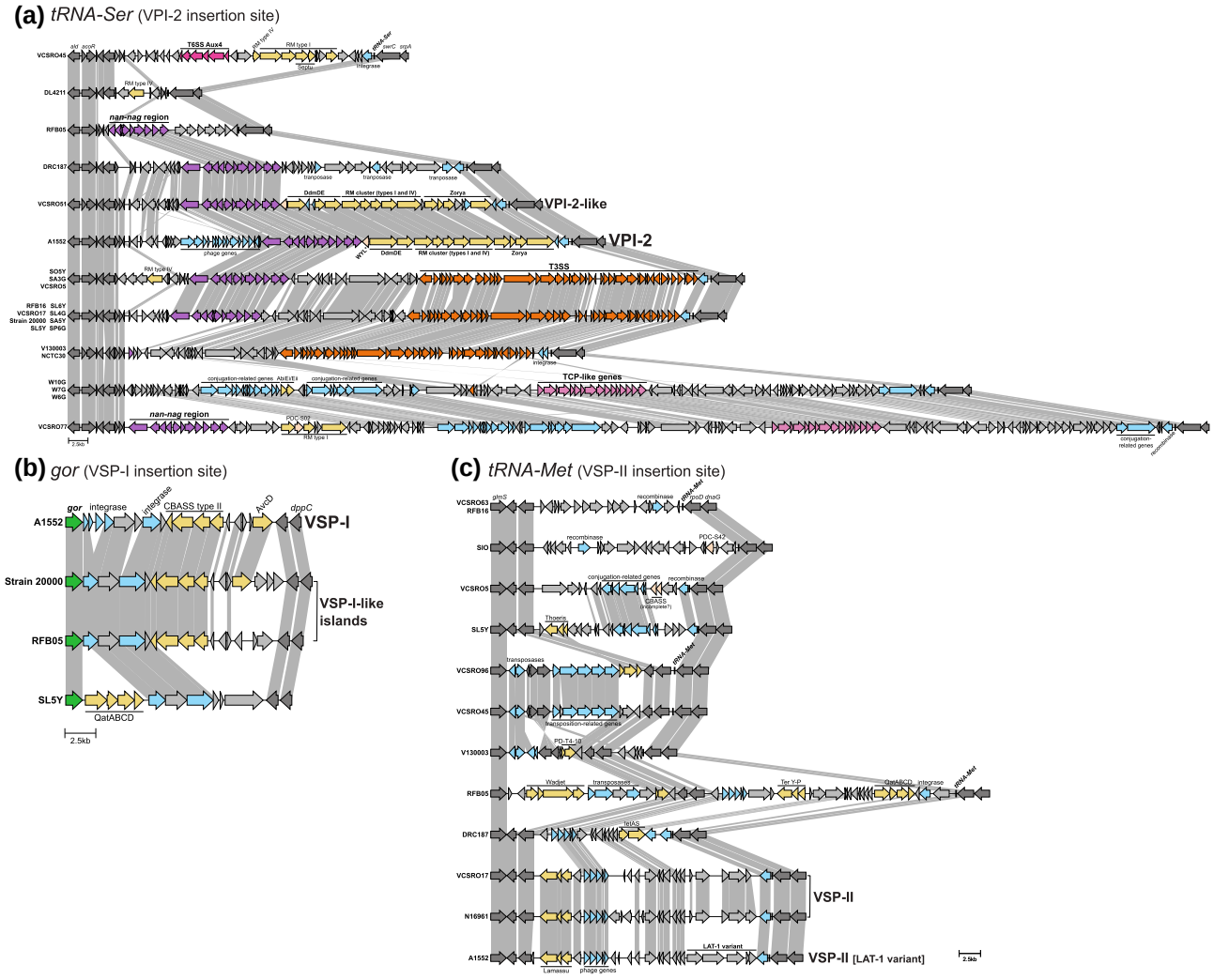

Fig. 4. Circulation of pandemic-like islands and gene clusters in nonpandemic *V. cholerae* MGEs. Comparative gene content within GIs located at the *tRNA-Ser* a), *gor* b), and *tRNA-Met* c) loci across various nonpandemic *V. cholerae* strains. Strains featuring these GIs are listed on the left. Light gray connections indicate protein identities above 30% (default parameter of the clinker pipeline). These connections are not present in the case of tRNA genes such as *tRNA-Ser* a) and *tRNA-Met* c), as comparisons in clinker are based on the encoded protein sequences. Insertion sites are marked in green (arrows in the case of protein-coding genes, small squares in the case of tRNA genes). Dark gray arrows point to genes flanking the islands, with important genes labeled. Light gray arrows denote genes encoding hypothetical proteins or those not directly related to the main findings. Light blue arrows highlight mobilization-related genes. Yellow arrows mark antiphage/antiplasmid defense genes or gene clusters, while Salmon-colored arrows indicate PDC, each labeled accordingly. The gene marked with "WYL*" was identified by a loose model that detects various defense-modification systems ("Dms_other") by PADLOC. The figure in a) depicts the *nan-nag* region from VPI-2 (in purple), T3SS gene clusters (in orange) and TCP-related (VPI-1-like) clusters (in pink) distributed in nonpandemic GIs. T6SS Aux4 cluster is indicated in dark pink. Each panel displays the pandemic island VPI-2 a), VSP-I b), and VSP-II c) of the 7PET strain A1552 for comparison.

A similar instance was observed for a VPI-2-like island in strain VCSRO51, an O51 serogroup strain from India (isolated in 1973) (Murase et al. 2022). This strain lacks the phage-related genes found in the *bona fide* VPI-2 (Figs. 3 and 4a). Additionally, the VPI-2-like island in VCSRO51 shows significant disruptions in genes associated with defense systems. Notably, the *ddmD* gene, a component of the DdmDE DNA defense module that targets plasmids for degradation (Jaskólska et al. 2022), is disrupted by two short transposase-related genes. Furthermore, within the predicted Zorya antiphage system gene cluster, *zorABCD*, there are gene insertions that encode for transposase-related proteins or proteins of unknown function (Fig. 4a). We also noticed a distribution of the sialic acid metabolism-related *nan-nag* region, an integral part of VPI-2 (Jermyn and Boyd 2005), in other nonpandemic strains, pointing to the spread of critical pandemic genetic components across different genomic backgrounds (Fig. 4a).

Further analysis identified nonpandemic strains with VSP-I-like islands, such as strain 20000, an O1

environmental isolate from Russia (2016), and the *V. tarriae* strain RFB05 from a lake in the USA (2017), both showcasing slight variations in their genomic makeup compared with VSP-I of pandemic strains (Figs. 3 and 4b), with the latter strain lacking the antiphage cytidine deaminase gene *avcD* (Hsueh et al. 2022). Lastly, strain SL5Y, isolated from the Californian coast in 2004 (Keymer et al. 2007), contains an island at the *gor* locus, similar to VSP-I and characterized by shared integrases. However, the gene content of this island is unique compared with other strains, notably including a gene cluster encoding a predicted QatABCD antiphage defense system among other distinctive features (Fig. 4b). These are the only occurrences of islands inserted at the *gor* site in our dataset.

We identified a VSP-II GI in strain VCSRO17, an O17 serogroup isolate from a 1968 patient in India, that is 99.7% identical to VSP-II from pandemic strains of the 1970s, such as N16961 (Figs. 3 and 4c). This observation underscores that MGEs can be shared between pandemic and nonpandemic strains of *V. cholerae*. Furthermore, in ten other nonpandemic strains, including *V. tarriae* RFB05, we identified islands located next to the *tRNA-Met* gene, the site in which VSP-II is integrated in 7PET strains (note that the annotation for this tRNA gene was retained as *tRNA-Met*, as initially proposed for the reference genome of strain N16961 (Heidelberg et al. 2000), despite predictions by the Prokaryotic Genome Annotation Pipeline (PGAP) (Tatusova et al. 2016) and tRNAscan-SE (Chan and Lowe 2019) software tools identifying the locus as *tRNA-Ile*). Some of these islands contain gene blocks similar to those found in pandemic VSP-II, including integrase and specific phage-like and hypothetical genes, as observed in strains DRC187 and RFB05 (Fig. 4c). However, the overall composition of these islands shows significant diversity.

Furthermore, islands found to be well-conserved among pandemic strains, though not recognized as "pandemic islands," were detected in nonpandemic strains. For instance, the Vc95 Retron found inserted next to *VC2384* in pandemic strains (Fig. 1b and supplementary fig. S2g, Supplementary Material online) was also detected in strains 2012Env-9 and 2012Env-390 (Fig. 3 and supplementary fig. S7, Supplementary Material online). Moreover, the exact island found next to *tdh* in chromosome 2 of pandemic strains (composed of hypothetical genes) was also found in nonpandemic strains W10G and VCSRO51, and a variation of this island was encountered in strain SO5Y (supplementary fig. S5, Supplementary Material online).

Remarkably, we did not find primary pandemic islands located outside their known positions (*ssrA*, *tRNA-Ser*, *gor*, and *tRNA-Met*), with the exception of an extra VSP-I copy found on chromosome 2 in two strains from the DRC, as previously noted (supplementary fig. S1e, Supplementary Material online). However, within our dataset, we identified four strains that possess TCP-like gene clusters in islands adjacent to *tRNA-Ser*, a location different from *ssrA* where the VPI-1 (or TCP) island is typically found in pandemic strains (Fig. 4a). Specifically, strain VCSRO77, isolated from a patient in India in 1976 (Murase et al. 2022), harbors a 135 kb-long island near *tRNA-Ser* that includes both a TCP-like cluster (as in VPI-1) and a *nan-nag* region (as in VPI-2) (Fig. 4a). This type of data was not captured in earlier PCR-based genetic screenings of *V. cholerae* samples collected worldwide, which had identified environmental strains carrying key genes like *tcpA* (encoding the major TCP pilin) and/or *toxT* (encoding the virulence regulator ToxT) (Rivera et al. 2001; Faruque et al. 2003, 2004). This finding within our relatively limited dataset highlights a scenario in which pandemic-like gene clusters coexist within the same GI, emphasizing (i) the potential role of nonpandemic strains as reservoirs for both virulence-associated and virulence-unrelated gene clusters and (ii) the possible emergence of hazardous combinations of disease-related gene clusters within MGEs.

Antiphage/Antiplasmid Defense Systems in MGEs

Bacteriophages are notably prevalent in marine environments, where they are estimated to cause 20% to 40% of bacterial mortality (Suttle 2007). The significant impact of phages on bacterial populations has been increasingly recognized, due to advances in computational biology and genome analysis (Roux et al. 2015). These developments have revealed the broad distribution of phages and highlighted the variety of antiphage defense mechanisms that bacteria possess (Hampton et al. 2020). Contrary to earlier beliefs that such defenses were limited mainly to CRISPR-Cas and restriction-modification (RM) systems, recent research has significantly broadened this view (Makarova et al. 2011; van Houte et al. 2016; Koonin et al. 2017; Doron et al. 2018; Bernheim and Sorek 2020; Hampton et al. 2020; Gao et al. 2020; Vassallo et al. 2022; Georjon and Bernheim 2023). It has revealed a vast array of antiphage defense strategies within bacterial genomes, now including hundreds of confirmed or predicted distinct mechanisms (DeWeirdt et al. 2025; Mordret et al. 2025). This broadening of known antiphage systems likely reflects the dynamic evolutionary arms race between bacteriophages and their bacterial hosts.

Antiphage defense systems in bacteria are primarily encoded within MGEs (Makarova et al. 2011; Benler et al. 2021; Hussain et al. 2021; LeGault et al. 2021; Piel et al. 2022; Rousset et al. 2022; Vassallo et al. 2022; Hochhauser et al. 2023; Patel and Maxwell 2023). This strategic positioning allows for the rapid exchange of defense mechanisms, essential in the bacterial-phage evolutionary arms race (Koonin et al. 2017; Bernheim and Sorek 2020; Hussain et al. 2021; Piel et al. 2022). In our investigation of nonpandemic *V. cholerae*, we explored the diversity

and abundance of defense systems, including both antiphage and antiplasmid mechanisms—collectively referred to as the "defensome" (Beavogui et al. 2024).

Defensome Diversity

In our analysis of nonpandemic *V. cholerae* strains, we quantified the prevalence of various antiphage defense systems, as depicted in Fig. 5a. Consistent with previous research (Tesson et al. 2022), RM systems emerged as the most prevalent defense mechanism, with 30 of the genomes (representing 83.3% of our dataset) containing at least one of these systems. Following closely, CRISPR-Cas systems were identified in 16 genomes, accounting for 47.2% of the strains we studied, despite their consistent absence from 7PET strains (Box et al. 2016). CRISPR-Cas Type I-F were particularly prevalent, found in islands inserted in *ssrA* and *ffs*, for example (supplementary figs. S4 and S8, Supplementary Material online). This prevalence aligns with the observations of McDonald and colleagues, who reported that Type I-F is the dominant CRISPR-Cas system within the Vibrionaceae family (McDonald et al. 2019), indicating a significant trend across a broad spectrum of species.

Other notable defense mechanisms identified include PD-λ-1, found in 10 genomes (27.7%), while Gabija and SoFic were detected in 9 genomes (25% of our dataset). Several defense systems were detected in nearly one fifth

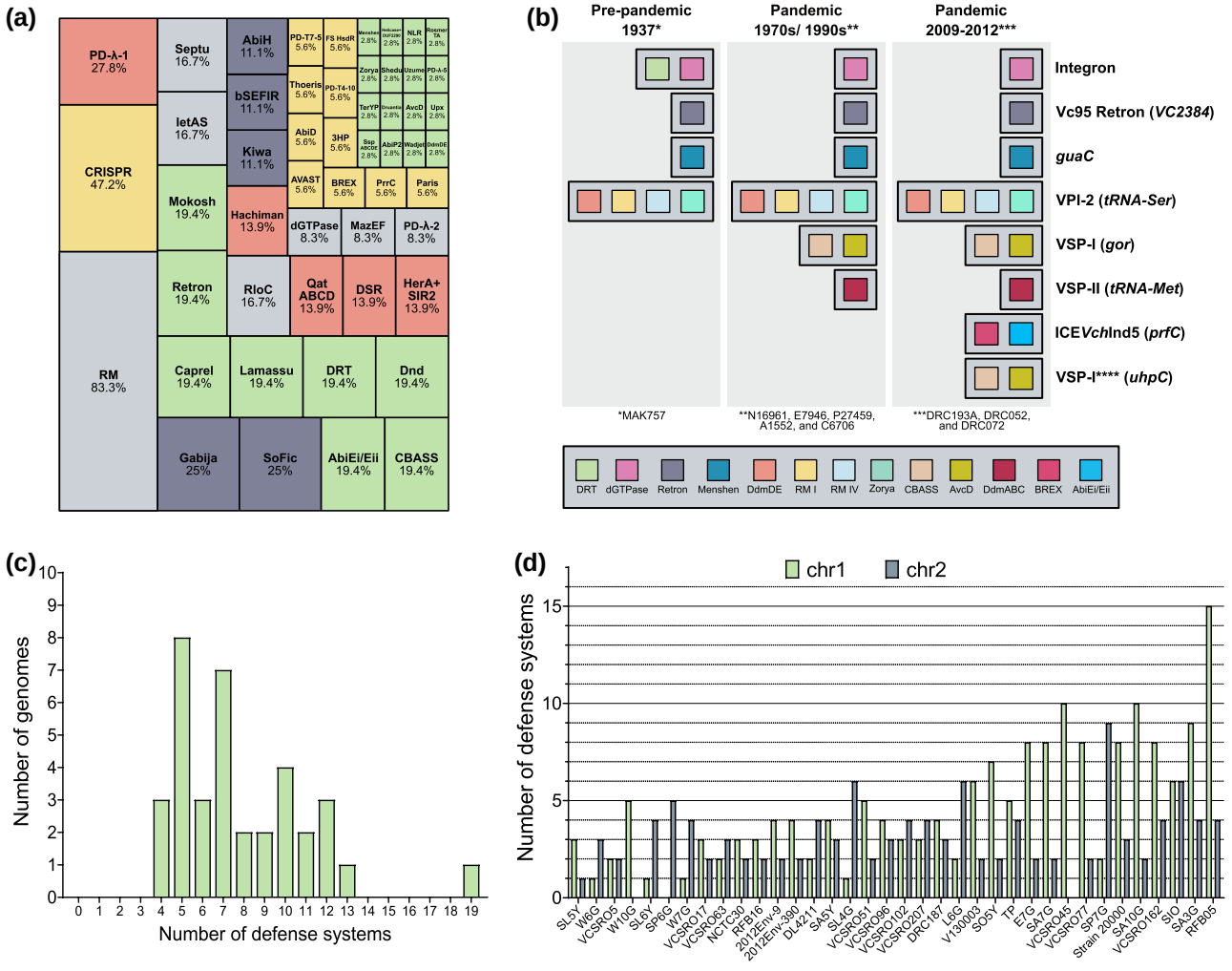

Fig. 5. Diversity and richness of defense systems in *V. cholerae*. a) Proportion of defense systems identified across genomes of nonpandemic *V. cholerae* in the current dataset. b) This figure represents defense systems found in the prepandemic strain MAK757 (left), pandemic strains isolated in the 1970s/1990s (center) and in 2009 to 2012 (right). Each square represents a defense system, following the color code depicted in the legend. Squares are included in rectangles representing the MGE in which they are found, which are indicated on the right of the figure. Specific strains of each isolation period are referenced at the bottom of each block, indicated by *, ** or ***. **** The second VSP-I copy inserted next to *uhpC* is only found in strains DRC052 and DRC072. c) Overview of the number of antiphage defense systems across nonpandemic *V. cholerae* strains, including the notable case of *V. tarriae* RFB05. d) Detailed comparison of defense system counts on chromosomes 1 and 2 across individual nonpandemic *V. cholerae* isolates.

of the strains (7 genomes or 19.4%): AbiEi/Eii, CBASS, CapRel, Lamassu, DRT, Dnd, Retron and Mokosh. IetAS, Septu, and RloC were identified in 6 genomes (16.6%). Additionally, we detected 36 distinct defense systems scattered across the genomes, though each was present in five or less strains (Fig. 5a). Importantly, using the newest version of PADLOC we detected a myriad of "phage defence candidates" (PDCs) and "Hma (for Helicase-Methylase-ATPase system)-embedded candidates" (Payne et al. 2024). These candidates were not considered as *bona fide* defense systems in our analyses, but their presence was incorporated in all figures throughout the manuscript. This varied distribution underscores the extensive diversity and specificity of antiphage/antiplasmid defense mechanisms within nonpandemic *V. cholerae*, highlighting the species' complex evolutionary adaptations to threats by phages and other MGEs.

Comparative Analysis of Defense Systems in Pandemic and Nonpandemic *V. cholerae* Strains

In pandemic *V. cholerae* strains, there is a noticeable conservation of defense systems across strains sharing identical GIs. For example, the prepandemic strain MAK757, dating back to 1937 and lacking VSP islands, contains seven predicted antiphage/antiplasmid systems (Fig. 5b). In contrast, strains from the 1970s and 1990s, such as N16961, E7946, P27459, A1552, and C6706, are equipped with nine identified defense systems (Fig. 5b). Note that the three recently reported defense systems of the West African–South American (WASA) lineage (Adams et al. 2025), which includes the Peruvian outbreak strains A1552 and C6706, are not included in this count, as neither DefenseFinder nor PADLOC was capable of detecting them at the time the programs were run. The most recently isolated strains from the DRC contain eleven (DRC193A) or thirteen (DRC052 and DRC072) systems (Fig. 5b). The latter group's additional defense systems are incorporated within their ICE (ICE*Vch*Ind5) located at the *prfC* locus, as previously demonstrated for such elements of Bangladeshi *V. cholerae* isolates (LeGault et al. 2021). DRC052 and DRC072 contain an additional VSP-I encoded next to *uhpC,* as previously mentioned (supplementary fig. S1e, Supplementary Material online, Fig. 5b). This pattern highlights the evolutionary conservation of antiphage/antiplasmid defenses in pandemic lineages and the introduction of novel defense systems through the incorporation of MGEs in the flexible genome.

In contrast, in our examination of nonpandemic *V. cholerae* strains, we discovered a broad spectrum of defense system counts per genome, ranging from 4 to 13 (with the exception of *V. tarriae* RFB05 strain, which possesses 19 systems) (Fig. 5c). The majority of strains in our study hosted approximately 5 antiphage/antiplasmid systems each, which is in line with previous surveys reporting an average of 5 defense systems per organism based on a vast dataset of 21,738 prokaryotic genomes (Tesson et al. 2022). We also observed a significant variation in the distribution of these defense systems across chromosomes (Fig. 5d). While certain strains localized all their defense mechanisms to either chromosome 1 or 2 (as seen in strains SP6G, W10G, and VCSRO45), others exhibited a more scattered arrangement, incorporating defense systems across both chromosomes without a clear pattern. There are several hypotheses that could lead to the disparity of defense system localization. While it is beyond the scope of this study to explore the defense system localization evolution, such variability could suggest a complex evolutionary landscape for the acquisition and distribution of antiphage/antiplasmid defense mechanisms in nonpandemic *V. cholerae* strains. In addition, it might reflect the adaptive strategies used against diverse phage challenges in the environments where these strains are found, but also chromosome-specific dynamics in defense systems acquisition.

Overall, *V. cholerae* showcases a wide array of defense mechanisms, positioning it among species with "many diverse systems," similar to *Pseudomonas aeruginosa* as noted by Tesson and colleagues (Tesson et al. 2022). Our research therefore highlights *V. cholerae*'s genomic diversity, particularly in its genes encoding antiphage/antiplasmid defense mechanisms. The observed diversity also underscores the importance of HGT in disseminating these defense systems.

Defensome Hotspots

In our subsequent analysis, we aimed to pinpoint potential genomic hotspots for the insertion of defense systems. To achieve this, we mapped the occurrence of GIs either possessing or lacking a defense system, across the insertion sites on chromosomes 1 and 2 (as illustrated in Fig. 6a and b). We identified several insertion sites that were frequently occupied across multiple strains but did not consistently contain islands with known defense systems. These include the O-antigen biosynthesis cluster and the sites next to *rtxA, pyrF*, and *katG* on chromosome 1 (Fig. 6a). It is important to note that the precision of our detection of defense systems heavily relied on the capabilities of the tools we used, namely DefenseFinder and PADLOC (Abby et al. 2014; Payne et al. 2021, 2022, 2024; Tesson et al. 2022). This reliance suggests the possibility that additional, yet unidentified, defense systems may be encoded within these islands, awaiting discovery with more targeted analytical tools and the incorporation of new defense systems into DefenseFinder and PADLOC (Abby et al. 2014; Payne et al. 2021, 2022, 2024; Tesson et al. 2022). Indeed, as mentioned above, we identified several candidate defense systems in our dataset using the most recent version of PADLOC (Payne et al. 2024), but did not include these in

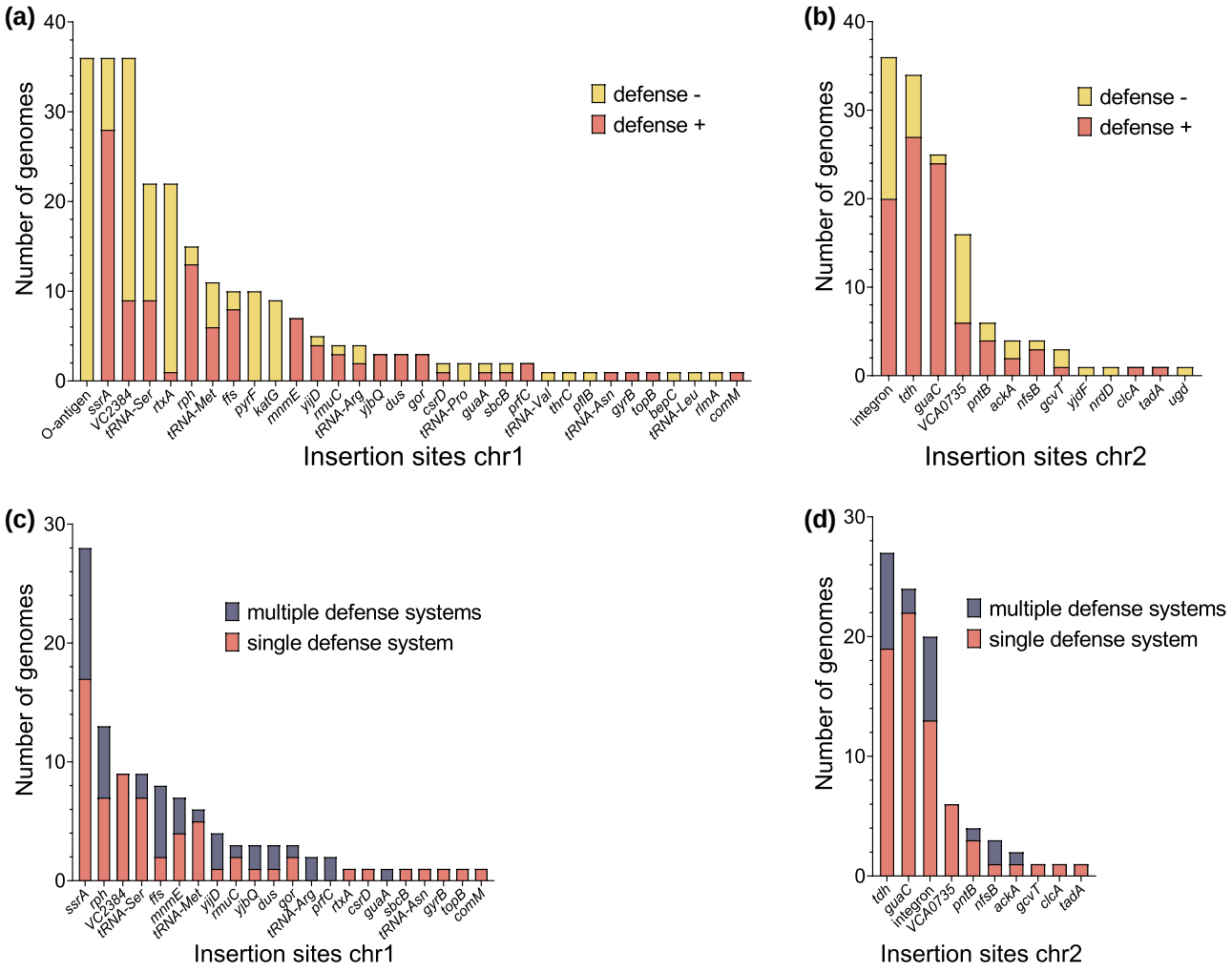

Fig. 6. Identifying genomic regions primed for defense system insertion and the presence of defense islands in nonpandemic *V. cholerae*. a, b) Visualization of MGE occupancy on chromosomes 1 (a) and 2 (b), distinguishing sites without defense systems (yellow) from those with one or more defense systems (red). c, d) Distribution of MGEs with defense systems across chromosomes 1 (c) and 2 (d). Each insertion site is plotted on the *x* axis, indicating the number of MGEs containing either a single (red) or multiple (gray) defense systems.

our analyses due to their provisional status. Importantly, the number of predicted defense systems would further increase with the use of AI-assisted defense prediction tools (DeWeirdt et al. 2025; Mordret et al. 2025).

We identified specific insertion sites that function as hotspots for defense systems, as illustrated in Fig. 6. Notable examples include *ssrA* (supplementary fig. S4, Supplementary Material online), *VC2384* (supplementary fig. S7, Supplementary Material online), *ffs, yjbQ* (supplementary fig. S8, Supplementary Material online), *rph,* and *mnmE* (supplementary fig. S9, Supplementary Material online) on chromosome 1, along with *tdh* and *guaC* on chromosome 2 (supplementary figs. S5 and S10, Supplementary Material online). Among these, the *guaC* insertion site was particularly prominent, with a majority of strains harboring short MGEs equipped with defense systems (supplementary fig. S10, Supplementary Material online). Similarly, short islands inserted next to *VC2384,* where the Vc95 Retron is located in 7PET, are frequently occupied by defense systems in nonpandemic strains (supplementary fig. S7, Supplementary Material online). In many cases, however, we detected genes coding for proteins of unknown functions or those bearing domains commonly linked to defense mechanisms, including ATPases and DNA-binding proteins (supplementary fig. S7, Supplementary Material online). Several of these genes were predicted to encode PDC by PADLOC (supplementary fig. S7, Supplementary Material online). Furthermore, our analysis revealed "hotspot regions" within GIs, consistently occupied by defense systems. For instance, the 3 ends of islands next to *tdh* (supplementary fig. S5, Supplementary Material online) and near

VCA0735 (supplementary fig. S11, Supplementary Material online) on chromosome 2 showcased such patterns. Within these regions, we also detected genes similar to those found in islands next to *VC2384*, such as helicases and ATPases. This leads us to hypothesize that these genes or gene clusters might represent defense systems that current detection tools, such as PADLOC and DefenseFinder, failed to recognize, as they might represent novel systems not yet characterized. A recent study on sedentary chromosomal integrons, including those in 7PET strain N16961, supports this idea (Darracq et al. 2024). Despite the absence of detected antiviral defense systems in these integron islands using DefenseFinder, experimental validation revealed that an astonishing 10% of integron cassettes exhibited antiphage activity when artificially expressed in *V. cholerae* or *E. coli*. This suggests that chromosomal integrons may serve as reservoirs for bacterial antiphage defense systems, a finding that has also been demonstrated for mobile integrons and the chromosomal integron of *V. parahaemolyticus* (Darracq et al. 2024; Kieffer et al. 2024; Getz et al. 2025).

Defense Islands

Defense systems in *V. cholerae* often cluster together, consistent with the concept of "defense islands" (Makarova et al. 2011; Doron et al. 2018; Hussain et al. 2021). Our interest was piqued by two main questions: (i) the prevalence of these defense islands within our dataset, and (ii) their frequency of occurrence at specific genomic sites. To explore these aspects, we cataloged all insertion sites that hosted at least one island equipped with a defense system and analyzed the occurrence of MGEs with one or more defense systems at each site. Our findings revealed that, across most insertion sites, islands with a singular predicted defense system predominated over those with multiple systems (Fig. 6c and d), although this observation is made within the limitations of our detection capabilities, as outlined above. However, certain locations, such as *ssrA* (supplementary fig. S4, Supplementary Material online), *ffs, yjbQ* (supplementary fig. S8, Supplementary Material online), *rph,* (supplementary fig. S9a, Supplementary Material online), *yijD* and *folD* (supplementary fig. S12a and b, Supplementary Material online), on chromosome 1, stood out as exceptions, where defense islands were more commonly found.

Prevalence of Prophage-Derived MGEs

We identified prophage signatures within our genomic dataset, encompassing both pandemic and nonpandemic strains, and determined whether these signatures coincided with the detected MGEs. We found that certain insertion sites frequently hosted MGEs bearing phage signatures. Notably, the *rtxA* site on chromosome 1, where the CTX prophage integrates in pandemic & toxigenic strains, was one such location (Fig. 3). Other sites, including *tRNA-Leu, tRNA-Pro, rlmA*, *pepN*, and *sbcB* on chromosome 1, and *ugd* on chromosome 2, were exclusively associated with prophage-related MGEs, although these sites were generally not frequently occupied (Fig. 3). Additionally, we identified specific strains where MGEs of prophage origin were present at particular sites, such as *tRNA-Ser, tRNA-Met,* and *tRNA-Arg* on chromosome 1 (Fig. 4a and c, and supplementary fig. S12b, Supplementary Material online), and *gcvT*, *tdh, pntB,* and *guaC* on chromosome 2 (supplementary figs. S3 and S5 and S6, and S10, Supplementary Material online). Interestingly, prophage signatures at the *tRNA-Ser* site were predominantly observed in VPI-2 from pandemic strains. However, ultimately, no clear correlation was found between the presence of prophage signatures and the existence of defense systems (Fig. 3).

Conclusion

To our knowledge, this study represents the first comprehensive examination of the mobilome and defensome diversity within the *V. cholerae* species, extending beyond the 7PET pandemic clade. This work establishes a foundational database cataloging MGEs and antiphage defense systems identified in these strains, highlighting their distribution, frequency, and variability. Although this study includes a limited number of nonpandemic strains, our findings point toward potentially new defense systems and unexplored molecular mechanisms. These discoveries offer rich opportunities for future scientific exploration and a deeper understanding of *V. cholerae* as a species.

Materials and Methods

Bacterial Strains and WGS

High-quality, well-assembled WGS of nonpandemic environmental or clinical *V. cholerae* strains were identified from existing literature. Details of the strains used are listed in supplementary tables S1 and S2, Supplementary Material online. Unannotated *V. cholerae* genomes were annotated using the PGAP (PGAP, 2022 to 2010-03. build6384) (Tatusova et al. 2016).

Strain MAK757 was genome sequenced using PacBio technology and de novo-assembled as described (Stutzmann and Blokesch 2020). The sequencing data have been deposited in NCBI (see data availability section).

Pangenome and Phylogenetic Reconstructions

The pangenome of all *V. cholerae* strains and *V. mimicus* was reconstructed using the PPanGGOLiN (v. 1.2.74) pipeline (all options, default parameters, sequences and annotation provided) (Gautreau et al. 2020). Identified core

genes were used for the subsequent phylogenetic reconstruction. To find the optimal model for each of the 1,532 core genes, ModelFinder (Kalyaanamoorthy et al. 2017) was used. The phylogeny of the 45 *V. cholerae* strains, with *V. mimicus* as an outgroup, was reconstructed using iqtree2 software (v2.2.0) (Minh et al. 2020). The model for each individual gene was implemented from ModelFinder, and 100 bootstraps were computed in iqtree2.

Detection of MGEs and Defense Systems

GI Identification

To detect MGEs, general patterns of genomic plasticity were analyzed using PanGGoLiN (Gautreau et al. 2020) based on the 45 genomes included in this study. RGPs (panRGP function, with default options) (Bazin et al. 2020) were thereby identified based on comparative genomics. Individual RGPs (i.e. GIs identified in each genome), were then validated using dedicated tools that predict GIs based on their sequence composition (see supplementary table S3, Supplementary Material online). All programs were run with default parameters. For each genome, five different GI searches were performed with different assumptions underlying each individual search (for more details see [da Silva Filho et al. 2018]). The results were subsequently visualized in Geneious Prime (v. 2022, 2023, and 2024; https://www.geneious.com) and manually curated. To be considered as a *bona fide* GI in the context of this study, the island had to fulfill several criteria: (i) to be detected by panRGP and (ii) confirmed by at least two among the other software tools used in this study (see supplementary table S3, Supplementary Material online); and (iii) to be at least 10 kb in length, as previously suggested (Juhas et al. 2009; Arnold et al. 2022). To have a broader overview of the defensome (Beavogui et al. 2024) diversity, an exception to this rule was made for islands encoding predicted antiphage/antiplasmid defense systems.

PGAP annotations were used as a first indication of the identity of gene clusters found in the islands. Additional BLAST searches (either for whole gene clusters or individual predicted proteins, using the online web service with default settings against the Reference proteins (refseq_protein) database) were used to identify gene clusters found in each GI, and extensive literature searches using PubMed were performed to characterize the identified clusters. To confirm the conservation (and therefore identity) of gene clusters among islands from different strains, we used clinker (v 0.0.25, default options) (Gilchrist and Chooi 2021). Details about this method are given below (*Alignment of GIs*).

Defense Systems Identification

To identify defense systems in the genomes that were included in our dataset, we used the tools DefenseFinder (accessed in April 2024) (Abby et al. 2014; Tesson et al. 2022) and PADLOC (v2.0.0) (Payne et al. 2021, 2022, 2024). Identified defense systems were visualized using Geneious Prime (v. 2024; https://www.geneious.com).

Prophage Identification

The online web tool PHASTER (Arndt et al. 2016) was used to identify putative prophage sequences within the genomes.

Alignment of GIs

For each insertion site in which GIs were detected, nucleotide sequences of the islands and neighboring genes were extracted. The comparison of the islands' gene content and organization was visualized using clinker (v 0.0.25, default options) (Gilchrist and Chooi 2021). Nucleotide sequences of islands from pandemic strains were extracted from the genomes and aligned using MAFFT (auto option, v.7508) (Katoh and Standley 2013). Each alignment was subsequently imported into Geneious Prime (v. 2022 and 2023; https://www.geneious.com) and distance matrices were calculated. Heatmaps of these matrices were generated in GraphPad Prism version 8.0.0 for MacOS (GraphPad Software, USA, www.graphpad.com).

Supplementary Material

Supplementary material is available at *Genome Biology and Evolution* online.

Acknowledgments

We thank the members of the Blokesch laboratory for fruitful discussions. We acknowledge the staff of the Lausanne Genomic Technologies Facility at the University of Lausanne and the members of the EPFL/UNIL Bioinformatics Competence Center for their contributions to sample processing, sequencing, genome assembly, and SMRT sequencing analysis. We acknowledge the use of ChatGPT for language editing, ensuring clarity and coherence, without the generation of new data. This work was supported by the Swiss National Science Foundation (310030_185022), a Consolidator Grant by the European Research Council (724630), and an International Research Scholarship by the Howard Hughes Medical Institute (HHMI) (grant 55008726) to M.B.

Author Contributions

N.C.D.D., A.L., and M.B. conceived the project and designed the details of the study; N.C.D.D. and A.L. developed the concept of the pipeline to detect GIs and defense systems; A.L. developed the pipeline and scripts used in this study; N.C.D.D. analyzed the data; M.B. oversaw the project's implementation, verified findings, and secured funding; N.C.D.D, A.L., and M.B. wrote the manuscript.

Data Availability

The whole-genome sequencing data of strain MAK757 have been deposited in NCBI under BioProject accession number PRJNA1028811. The genome sequence is available in GenBank under accession numbers CP159790 and CP159791 for chromosome 1 and chromosome 2, respectively. The raw reads are available from the Sequence Read Archive under submission numbers SRS21856092.

Literature Cited

Abby SS, Néron B, Ménager H, Touchon M, Rocha EP. Macsyfinder: a program to mine genomes for molecular systems with an application to CRISPR-Cas systems. PLoS One. 2014:9(10):e110726. https://doi.org/10.1371/journal.pone.0110726.

Adams DW, Jaskólska M, Lemopoulos A, Stutzmann S, Righi L, Bader L, Blokesch M. West African–South American pandemic *Vibrio cholerae* encodes multiple distinct phage defence systems. Nat Microbiol. 2025:in press. https://doi.org/10.1038/s41564-025-02004-9.

Ahmed N, Dobrindt U, Hacker J, Hasnain SE. Genomic fluidity and pathogenic bacteria: applications in diagnostics, epidemiology and intervention. Nat Rev Microbiol. 2008:6(5):387–394. https://doi.org/10.1038/nrmicro1889.

Altindis E, Dong T, Catalano C, Mekalanos J. Secretome analysis of *Vibrio cholerae* type VI secretion system reveals a new effector-immunity pair. mBio. 2015:6(2):e00075-15. https://doi.org/10.1128/mBio.00075-15.

Arndt D, Grant JR, Marcu A, Sajed T, Pon A, Liang Y, Wishart DS. PHASTER: a better, faster version of the PHAST phage search tool. Nucleic Acids Res. 2016:44(W1):W16–W21. https://doi.org/10.1093/nar/gkw387.

Arnold BJ, Huang IT, Hanage WP. Horizontal gene transfer and adaptive evolution in bacteria. Nat Rev Microbiol. 2022:20(4):206–218. https://doi.org/10.1038/s41579-021-00650-4.

Azarian T, Ali A, Johnson JA, Jubair M, Cella E, Ciccozzi M, Nolan DJ, Farmerie W, Rashid MH, Sinha-Ray S, et al. Non-toxigenic environmental *Vibrio cholerae* O1 strain from Haiti provides evidence of pre-pandemic cholera in Hispaniola. Sci Rep. 2016:6:36115. https://doi.org/10.1038/srep36115.

Azarian T, Ali A, Johnson JA, Mohr D, Prosperi M, Veras NM, Jubair M, Strickland SL, Rashid MH, Alam MT, et al. Phylodynamic analysis of clinical and environmental *Vibrio cholerae* isolates from Haiti reveals diversification driven by positive selection. mBio. 2014:5(6):e01824-14. https://doi.org/10.1128/mBio.01824-14.

Bazin A, Gautreau G, Médigue C, Vallenet D, Calteau A. panRGP: a pangenome-based method to predict genomic islands and explore their diversity. Bioinformatics. 2020:36(Suppl 2):i651–i658. https://doi.org/10.1093/bioinformatics/btaa792.

Beavogui A, Lacroix A, Wiart N, Poulain J, Delmont TO, Paoli L, Wincker P, Oliveira PH. The defensome of complex bacterial communities. Nat Commun. 2024:15(1):2146. https://doi.org/10.1038/s41467-024-46489-0.

Benler S, Faure G, Altae-Tran H, Shmakov S, Zheng F, Koonin E. Cargo genes of Tn7-like transposons comprise an enormous diversity of defense systems, mobile genetic elements, and antibiotic resistance genes. mBio. 2021:12(6):e02938-21. https://doi.org/10.1128/mBio.02938-21.

Bernardy EE, Turnsek MA, Wilson SK, Tarr CL, Hammer BK. 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. 2016:82(9):2833–2842. https://doi.org/10.1128/AEM.00351-16.

Bernheim A, Sorek R. The pan-immune system of bacteria: antiviral defence as a community resource. Nat Rev Microbiol. 2020:18(2):113–119. https://doi.org/10.1038/s41579-019-0278-2.

Bobonis J, Mitosch K, Mateus A, Karcher N, Kritikos G, Selkrig J, Zietek M, Monzon V, Pfalz B, Garcia-Santamarina S, et al. Bacterial retrons encode phage-defending tripartite toxin-antitoxin systems. Nature. 2022:609(7925):144–150. https://doi.org/10.1038/s41586-022-05091-4.

Boucher Y. Sustained local diversity of *Vibrio cholerae* O1 biotypes in a previously cholera-free country. mBio. 2016:7(3):e00570-16. https://doi.org/10.1128/mBio.00570-16.

Box AM, McGuffie MJ, O'Hara BJ, Seed KD. Functional analysis of bacteriophage immunity through a type I-E CRISPR-cas system in *Vibrio cholerae* and its application in bacteriophage genome engineering. J Bacteriol. 2016:198(3):578–590. https://doi.org/10.1128/JB.00747-15.

Boyd EF, Almagro-Moreno S, Parent MA. Genomic islands are dynamic, ancient integrative elements in bacterial evolution. Trends Microbiol. 2009:17(2):47–53. https://doi.org/10.1016/j.tim.2008.11.003.

Boyd EF, Carpenter MR, Chowdhury N, Cohen AL, Haines-Menges BL, Kalburge SS, Kingston JJ, Lubin JB, Ongagna-Yhombi SY, Whitaker WB. Post-genomic analysis of members of the family *Vibrionaceae*. Microbiol Spectr. 2015:3. https://doi.org/10.1128/microbiolspec.VE-0009-2014.

Boyer F, Fichant G, Berthod J, Vandenbrouck Y, Attree I. Dissecting the bacterial type VI secretion system by a genome wide in silico analysis: what can be learned from available microbial genomic resources? BMC Genomics. 2009:10(1):104. https://doi.org/10.1186/1471-2164-10-104.

Büttner D. Protein export according to schedule: architecture, assembly, and regulation of type III secretion systems from plant- and animal-pathogenic bacteria. Microbiol Mol Biol Rev. 2012:76(2):262–310. https://doi.org/10.1128/MMBR.05017-11.

Carpenter MR, Kalburge SS, Borowski JD, Peters MC, Colwell RR, Boyd EF. CRISPR-Cas and contact-dependent secretion systems present on excisable pathogenicity islands with conserved recombination modules. J Bacteriol. 2017:199(10):e00842-16. https://doi.org/10.1128/JB.00842-16.

Chan PP, Lowe TM. tRNAscan-SE: searching for tRNA genes in genomic sequences. Methods Mol Biol. 2019:1962:1–14. https://doi.org/10.1007/978-1-4939-9173-0_1.

Chattopadhyay DJ, Sarkar BL, Ansari MQ, Chakrabarti BK, Roy MK, Ghosh AN, Pal SC. New phage typing scheme for *Vibrio cholerae* O1 biotype El Tor strains. J Clin Microbiol. 1993:31(6):1579–1585. https://doi.org/10.1128/jcm.31.6.1579-1585.1993.

Chun J, Grim CJ, Hasan NA, Lee JH, Choi SY, Haley BJ, Taviani E, Jeon YS, Kim DW, Lee JH, et al. Comparative genomics reveals mechanism for short-term and long-term clonal transitions in pandemic *Vibrio cholerae*. Proc Natl Acad Sci U S A. 2009:106(36):15442–15447. https://doi.org/10.1073/pnas.0907787106.

Clemens JD, van Loon F, Sack DA, Rao MR, Ahmed F, Chakrabort YJ, Kay BA, Khan MR, Yunus MD, Harris JR, et al. Biotype as determinant of natural immunising effect of cholera. Lancet. 1991:337(8746):883–884. https://doi.org/10.1016/0140-6736(91)90207-6.

Costa TR, Felisberto-Rodrigues C, Meir A, Prevost MS, Redzej A, Trokter M, Waksman G. Secretion systems in gram-negative bacteria: structural and mechanistic insights. Nat Rev Microbiol. 2015:13(6):343–359. https://doi.org/10.1038/nrmicro3456.

Cottingham KL, Chiavelli DA, Taylor RK. Environmental microbe and human pathogen: the ecology and microbiology of *Vibrio*

cholerae. Front Ecol Environ Microbiol. 2003:1(2):80–86. https://doi.org/10.1890/1540-9295(2003)001[0080:EMAHPT]2.0.CO;2.

Crisan CV, Chande AT, Williams K, Raghuram V, Rishishwar L, Steinbach G, Watve SS, Yunker P, Jordan IK, Hammer BK. Analysis of *Vibrio cholerae* genomes identifies new type VI secretion system gene clusters. Genome Biol. 2019:20(1):163. https://doi.org/10.1186/s13059-019-1765-5.

Crisan CV, Hammer BK. The *Vibrio cholerae* type VI secretion system: toxins, regulators and consequences. Environ Microbiol. 2020:22(10):4112–4122. https://doi.org/10.1111/1462-2920.14976.

Darracq B, Littner E, Brunie M, Bos J, Kaminski P-A, Depardieu F, Slesak W, Debatisse K, Touchon M, Bernheim A, et al. Sedentary chromosomal integrons as biobanks of bacterial anti-phage defence systems. bioRxiv 601686. https://doi.org/10.1101/2024.07.02.601686, 3 July 2024, preprint: not peer reviewed.

da Silva Filho AC, Raittz RT, Guizelini D, De Pierri CR, Augusto DW, Dos Santos-Weiss ICR, Marchaukoski JN. Comparative analysis of genomic island prediction tools. Front Genet. 2018:9:619. https://doi.org/10.3389/fgene.2018.00619.

Davies BW, Bogard RW, Young TS, Mekalanos JJ. Coordinated regulation of accessory genetic elements produces cyclic di-nucleotides for *V. cholerae* virulence. Cell. 2012:149(2):358–370. https://doi.org/10.1016/j.cell.2012.01.053.

DeWeirdt PC, Mahoney EM, Laub MT. DefensePredictor: a machine learning model to discover novel prokaryotic immune systems. bioRxiv 631726. https://doi.org/10.1101/2025.01.08.631726, 8 January 2025, preprint: not peer reviewed.

Dobrindt U, Hochhut B, Hentschel U, Hacker J. Genomic islands in pathogenic and environmental microorganisms. Nat Rev Microbiol. 2004:2(5):414–424. https://doi.org/10.1038/nrmicro884.

Doolittle WF, Papke RT. Genomics and the bacterial species problem. Genome Biol. 2006:7(9):116. https://doi.org/10.1186/gb-2006-7-9-116.

Dorman MJ, Kane L, Domman D, Turnbull JD, Cormie C, Fazal MA, Goulding DA, Russell JE, Alexander S, Thomson NR. The history, genome and biology of NCTC 30: a non-pandemic *Vibrio cholerae* isolate from world war one. Proc Biol Sci. 2019:286(1900):20182025. https://doi.org/10.1098/rspb.2018.2025.

Doron S, Melamed S, Ofir G, Leavitt A, Lopatina A, Keren M, Amitai G, Sorek R. Systematic discovery of antiphage defense systems in the microbial pangenome. Science. 2018:359(6379):eaar4120. https://doi.org/10.1126/science.aar4120.

Dziejman M, Balon E, Boyd D, Fraser CM, Heidelberg JF, Mekalanos JJ. Comparative genomic analysis of *Vibrio cholerae*: genes that correlate with cholera endemic and pandemic disease. Proc Natl Acad Sci U S A. 2002:99(3):1556–1561. https://doi.org/10.1073/pnas.042667999.

Dziejman M, Serruto D, Tam VC, Sturtevant D, Diraphat P, Faruque SM, Rahman MH, Heidelberg JF, Decker J, Li L, et al. Genomic characterization of non-O1, non-O139 *Vibrio cholerae* reveals genes for a type III secretion system. Proc Natl Acad Sci U S A. 2005:102(9):3465–3470. https://doi.org/10.1073/pnas.0409918102.

Faruque SM, Albert MJ, Mekalanos JJ. Epidemiology, genetics, and ecology of toxigenic *Vibrio cholerae*. Microbiol Mol Biol Rev. 1998:62(4):1301–1314. https://doi.org/10.1128/MMBR.62.4.1301-1314.1998.

Faruque SM, Chowdhury N, Kamruzzaman M, Dziejman M, Rahman MH, Sack DA, Nair GB, Mekalanos JJ. Genetic diversity and virulence potential of environmental *Vibrio cholerae* population in a cholera-endemic area. Proc Natl Acad Sci U S A. 2004:101(7):2123–2128. https://doi.org/10.1073/pnas.0308485100.

Faruque SM, Kamruzzaman M, Meraj IM, Chowdhury N, Nair GB, Sack RB, Colwell RR, Sack DA. Pathogenic potential of environmental *Vibrio cholerae* strains carrying genetic variants of the toxin-coregulated pilus pathogenicity island. Infect Immun. 2003:71(2):1020–1025. https://doi.org/10.1128/IAI.71.2.1020-1025.2003.

Frost LS, Leplae R, Summers AO, Toussaint A. Mobile genetic elements: the agents of open source evolution. Nat Rev Microbiol. 2005:3(9):722–732. https://doi.org/10.1038/nrmicro1235.

Galán JE, Waksman G. Protein-injection machines in bacteria. Cell. 2018:172(6):1306–1318. https://doi.org/10.1016/j.cell.2018.01.034.

Gao L, Altae-Tran H, Böhning F, Makarova KS, Segel M, Schmid-Burgk JL, Koob J, Wolf YI, Koonin EV, Zhang F. Diverse enzymatic activities mediate antiviral immunity in prokaryotes. Science. 2020:369(6507):1077–1084. https://doi.org/10.1126/science.aba0372.

Garza DR, Thompson CC, Loureiro EC, Dutilh BE, Inada DT, Junior EC, Cardoso JF, Nunes MR, de Lima CP, Silvestre RV, et al. Genome-wide study of the defective sucrose fermenter strain of *Vibrio cholerae* from the Latin American cholera epidemic. PLoS One. 2012:7(5):e37283. https://doi.org/10.1371/journal.pone.0037283.

Gautreau G, Bazin A, Gachet M, Planel R, Burlot L, Dubois M, Perrin A, Médigue C, Calteau A, Cruveiller S, et al. PPanGGOLiN: depicting microbial diversity via a partitioned pangenome graph. PLoS Comput Biol. 2020:16(3):e1007732. https://doi.org/10.1371/journal.pcbi.1007732.

Gennari M, Ghidini V, Caburlotto G, Lleo MM. Virulence genes and pathogenicity islands in environmental *Vibrio* strains nonpathogenic to humans. FEMS Microbiol Ecol. 2012:82(3):563–573. https://doi.org/10.1111/j.1574-6941.2012.01427.x.

Georjon H, Bernheim A. The highly diverse antiphage defence systems of bacteria. Nat Rev Microbiol. 2023:21(10):686–700. https://doi.org/10.1038/s41579-023-00934-x.

Getz LJ, Fairburn SR, Vivian Liu Y, Qian AL, Maxwell KL. Integrons are anti-phage defence libraries in *Vibrio parahaemolyticus*. Nat Microbiol. 2025:10(3):724–733. https://doi.org/10.1038/s41564-025-01927-7.

Gilchrist CLM, Chooi YH. Clinker & clustermap.js: automatic generation of gene cluster comparison figures. Bioinformatics. 2021:37(16):2473–2475. https://doi.org/10.1093/bioinformatics/btab007.

Gomez JB, Waters CM. A *Vibrio cholerae* type IV restriction system targets glucosylated 5-hydroxymethylcytosine to protect against phage infection. J Bacteriol. 2024:206(9):e0014324. https://doi.org/10.1128/jb.00143-24.

Hampton HG, Watson BNJ, Fineran PC. The arms race between bacteria and their phage foes. Nature. 2020:577(7790):327–336. https://doi.org/10.1038/s41586-019-1894-8.

Heidelberg JF, Eisen JA, Nelson WC, Clayton RA, Gwinn ML, Dodson RJ, Haft DH, Hickey EK, Peterson JD, Umayam L, et al. DNA sequence of both chromosomes of the cholera pathogen *Vibrio cholerae*. Nature. 2000:406(6795):477–483. https://doi.org/10.1038/35020000.

Hersch SJ, Watanabe N, Stietz MS, Manera K, Kamal F, Burkinshaw B, Lam L, Pun A, Li M, Savchenko A, et al. Envelope stress responses defend against type six secretion system attacks independently of immunity proteins. Nat Microbiol. 2020:5(5):706–714. https://doi.org/10.1038/s41564-020-0672-6.

Hochhauser D, Millman A, Sorek R. The defense island repertoire of the *Escherichia coli* pan-genome. PLoS Genet. 2023:19(4):e1010694. https://doi.org/10.1371/journal.pgen.1010694.

Hochhut B, Waldor MK. Site-specific integration of the conjugal *Vibrio cholerae* SXT element into prfC. Mol Microbiol. 1999:32(1):99–110. https://doi.org/10.1046/j.1365-2958.1999.01330.x.

Hsueh BY, Severin GB, Elg CA, Waldron EJ, Kant A, Wessel AJ, Dover JA, Rhoades CR, Ridenhour BJ, Parent KN, et al. Phage defence by deaminase-mediated depletion of deoxynucleotides in bacteria. Nat Microbiol. 2022:7(8):1210–1220. https://doi.org/10.1038/s41564-022-01162-4.

Hussain FA, Dubert J, Elsherbini J, Murphy M, VanInsberghe D, Arevalo P, Kauffman K, Rodino-Janeiro BK, Gavin H, Gomez A, et al. Rapid evolutionary turnover of mobile genetic elements drives bacterial resistance to phages. Science. 2021:374(6566):488–492. https://doi.org/10.1126/science.abb1083.

Innamorati KA, Earl JP, Aggarwal SD, Ehrlich GD, Hiller NL. The bacterial guide to designing a diversified gene portfolio. In: Tettelin H, Medini D, editors. The pangenome: diversity, dynamics and evolution of genomes. Cham (CH): Springer; 2020. p. 51–87.

Islam MT, Liang K, Orata FD, Im MS, Alam M, Lee CC, Boucher YF. *Vibrio tarriae* sp. nov., a novel member of the Cholerae clade. Int J Syst Evol Microbiol. 2022:72:005571. https://doi.org/10.1099/ijsem.0.005571.

Jaskólska M, Adams DW, Blokesch M. Two defence systems eliminate plasmids from seventh pandemic *Vibrio cholerae*. Nature. 2022:604(7905):323–329. https://doi.org/10.1038/s41586-022-04546-y.

Jermyn WS, Boyd EF. Molecular evolution of *Vibrio* pathogenicity island-2 (VPI-2): mosaic structure among *Vibrio cholerae* and *Vibrio mimicus* natural isolates. Microbiology. 2005:151(1):311–322. https://doi.org/10.1099/mic.0.27621-0.

Juhas M, van der Meer JR, Gaillard M, Harding RM, Hood DW, Crook DW. Genomic islands: tools of bacterial horizontal gene transfer and evolution. FEMS Microbiol Rev. 2009:33(2):376–393. https://doi.org/10.1111/j.1574-6976.2008.00136.x.

Kalyaanamoorthy S, Minh BQ, Wong TKF, von Haeseler A, Jermiin LS. ModelFinder: fast model selection for accurate phylogenetic estimates. Nat Methods. 2017:14(6):587–589. https://doi.org/10.1038/nmeth.4285.

Karaolis DK, Johnson JA, Bailey CC, Boedeker EC, Kaper JB, Reeves PR. A *Vibrio cholerae* pathogenicity island associated with epidemic and pandemic strains. Proc Natl Acad Sci U S A. 1998:95(6):3134–3139. https://doi.org/10.1073/pnas.95.6.3134.

Karaolis DK, Lan R, Reeves PR. The sixth and seventh cholera pandemics are due to independent clones separately derived from environmental, nontoxigenic, non-O1 *Vibrio cholerae*. J Bacteriol. 1995:177(11):3191–3198. https://doi.org/10.1128/jb.177.11.3191-3198.1995.

Katoh K, Standley DM. MAFFT multiple sequence alignment software version 7: improvements in performance and usability. Mol Biol Evol. 2013:30(4):772–780. https://doi.org/10.1093/molbev/mst010.

Keymer DP, Miller MC, Schoolnik GK, Boehm AB. Genomic and phenotypic diversity of coastal *Vibrio cholerae* strains is linked to environmental factors. Appl Environ Microbiol. 2007:73(11):3705–3714. https://doi.org/10.1128/AEM.02736-06.

Kieffer N, Hipólito A, Ortiz-Miravalles L, Blanco P, Delobelle T, Vizuete P, Ojeda FM, Jové T, Jurenas D, García-Quintanilla M, et al. Mobile integrons encode phage defense systems. bioRxiv 601719. https://doi.org/10.1101/2024.07.02.601719, 26 July 2024, preprint: not peer reviewed.

Koonin EV, Makarova KS, Wolf YI. Evolutionary genomics of defense systems in archaea and bacteria. Annu Rev Microbiol. 2017:71(1):233–261. https://doi.org/10.1146/annurev-micro-090816-093830.

Labbate M, Orata FD, Petty NK, Jayatilleke ND, King WL, Kirchberger PC, Allen C, Mann G, Mutreja A, Thomson NR, et al. A genomic island in *Vibrio cholerae* with VPI-1 site-specific recombination characteristics contains CRISPR-Cas and type VI secretion modules. Sci Rep. 2016:6:36891. https://doi.org/10.1038/srep36891.

LeGault KN, Hays SG, Angermeyer A, McKitterick AC, Johura FT, Sultana M, Ahmed T, Alam M, Seed KD. Temporal shifts in antibiotic resistance elements govern phage-pathogen conflicts. Science. 2021:373(6554):eabg2166. https://doi.org/10.1126/science.abg2166.

Lemopoulos A, Miwanda B, Drebes Dörr NC, Stutzmann S, Bompangue D, Muyembe-Tamfum J-J, Blokesch M. Genome sequences of *Vibrio cholerae* strains isolated in the DRC between 2009 and 2012. Microbiol Resour Announc. 2024:13(3):e0082723. https://doi.org/10.1128/mra.00827-23.

Makarova KS, Wolf YI, Snir S, Koonin EV. Defense islands in bacterial and archaeal genomes and prediction of novel defense systems. J Bacteriol. 2011:193(21):6039–6056. https://doi.org/10.1128/JB.05535-11.

Matthey N, Drebes Dörr NC, Blokesch M. Long-read-based genome sequences of pandemic and environmental *Vibrio cholerae* strains. Microbiol Resour Announc. 2018:7(23):e01574-18. https://doi.org/10.1128/mra.01574-18.

McDonald ND, Regmi A, Morreale DP, Borowski JD, Boyd EF. CRISPR-Cas systems are present predominantly on mobile genetic elements in *Vibrio* species. BMC Genomics. 2019:20(1):105. https://doi.org/10.1186/s12864-019-5439-1.

Miller KA, Tomberlin KF, Dziejman M. *Vibrio* variations on a type three theme. Curr Opin Microbiol. 2019:47:66–73. https://doi.org/10.1016/j.mib.2018.12.001.

Miller MC, Keymer DP, Avelar A, Boehm AB, Schoolnik GK. Detection and transformation of genome segments that differ within a coastal population of *Vibrio cholerae* strains. Appl Environ Microbiol. 2007:73(11):3695–3704. https://doi.org/10.1128/AEM.02735-06.

Miller VL, DiRita VJ, Mekalanos JJ. Identification of *toxS*, a regulatory gene whose product enhances ToxR-mediated activation of the cholera toxin promoter. J Bacteriol. 1989:171(3):1288–1293. https://doi.org/10.1128/jb.171.3.1288-1293.1989.

Millman A, Bernheim A, Stokar-Avihail A, Fedorenko T, Voichek M, Leavitt A, Oppenheimer-Shaanan Y, Sorek R. Bacterial retrons function in anti-phage defense. Cell. 2020:183(6):1551–-1561.e12. https://doi.org/10.1016/j.cell.2020.09.065.

Minh BQ, Schmidt HA, Chernomor O, Schrempf D, Woodhams MD, von Haeseler A, Lanfear R. IQ-TREE 2: new models and efficient methods for phylogenetic inference in the genomic era. Mol Biol Evol. 2020:37(5):1530–1534. https://doi.org/10.1093/molbev/msaa015.

Mordret E, Hervé A, Vaysset H, Clabby T, Tesson F, Shomar H, Lavenir R, Cury J, Bernheim A. Protein and genomic language models chart a vast landscape of antiphage defenses. bioRxiv 631966. https://doi.org/10.1101/2025.01.08.631966, 8 January 2025, preprint: not peer reviewed.

Murase K, Arakawa E, Izumiya H, Iguchi A, Takemura T, Kikuchi T, Nakagawa I, Thomson NR, Ohnishi M, Morita M. Genomic dissection of the *Vibrio cholerae* O-serogroup global reference strains: reassessing our view of diversity and plasticity between two chromosomes. Microb Genom. 2022:8(8):mgen000860. https://doi.org/10.1099/mgen.0.000860.

Murphy RA, Boyd EF. Three pathogenicity islands of *Vibrio cholerae* can excise from the chromosome and form circular intermediates. J Bacteriol. 2008:190(2):636–647. https://doi.org/10.1128/JB.00562-07.

Mutreja A, Kim DW, Thomson NR, Connor TR, Lee JH, Kariuki S, Croucher NJ, Choi SY, Harris SR, Lebens M, et al. Evidence for several waves of global transmission in the seventh cholera pandemic. Nature. 2011:477(7365):462–465. https://doi.org/10.1038/nature10392.

Ochman H, Lawrence JG, Groisman EA. Lateral gene transfer and the nature of bacterial innovation. Nature. 2000:405(6784):299–304. https://doi.org/10.1038/35012500.

O'Shea YA, Reen FJ, Quirke AM, Boyd EF. Evolutionary genetic analysis of the emergence of epidemic *Vibrio cholerae* isolates on the basis

of comparative nucleotide sequence analysis and multilocus virulence gene profiles. J Clin Microbiol. 2004:42(10):4657–4671. https://doi.org/10.1128/JCM.42.10.4657-4671.2004.

Otto SB, Servajean R, Lemopoulos A, Bitbol AF, Blokesch M. Interactions between pili affect the outcome of bacterial competition driven by the type VI secretion system. Curr Biol. 2024:34(11): 2403–2417.e9. https://doi.org/10.1016/j.cub.2024.04.041.

Patel PH, Maxwell KL. Prophages provide a rich source of antiphage defense systems. Curr Opin Microbiol. 2023:73:102321. https://doi.org/10.1016/j.mib.2023.102321.

Payne LJ, Hughes TCD, Fineran PC, Jackson SA. New antiviral defences are genetically embedded within prokaryotic immune systems. bioRxiv 577857. https://doi.org/10.1101/2024.01.29.577857, 30 January 2024, preprint: not peer reviewed.

Payne LJ, Meaden S, Mestre MR, Palmer C, Toro N, Fineran PC, Jackson SA. PADLOC: a web server for the identification of antiviral defence systems in microbial genomes. Nucleic Acids Res. 2022:50(W1): W541–W550. https://doi.org/10.1093/nar/gkac400.

Payne LJ, Todeschini TC, Wu Y, Perry BJ, Ronson CW, Fineran PC, Nobrega FL, Jackson SA. Identification and classification of antiviral defence systems in bacteria and archaea with PADLOC reveals new system types. Nucleic Acids Res. 2021:49(19):10868–10878. https://doi.org/10.1093/nar/gkab883.

Pearson GD, Woods A, Chiang SL, Mekalanos JJ. CTX genetic element encodes a site-specific recombination system and an intestinal colonization factor. Proc Natl Acad Sci U S A. 1993:90(8):3750–3754. https://doi.org/10.1073/pnas.90.8.3750.

Piel D, Bruto M, Labreuche Y, Blanquart F, Goudenège D, Barcia-Cruz R, Chenivesse S, Le Panse S, James A, Dubert J, et al. Phage-host coevolution in natural populations. Nat Microbiol. 2022:7(7): 1075–1086. https://doi.org/10.1038/s41564-022-01157-1.

Pukatzki S, Ma AT, Sturtevant D, Krastins B, Sarracino D, Nelson WC, Heidelberg JF, Mekalanos JJ. Identification of a conserved bacterial protein secretion system in *Vibrio cholerae* using the *Dictyostelium* host model system. Proc Natl Acad Sci U S A. 2006:103(5): 1528–1533. https://doi.org/10.1073/pnas.0510322103.

Rapa RA, Islam A, Monahan LG, Mutreja A, Thomson N, Charles IG, Stokes HW, Labbate M. A genomic island integrated into *recA* of *Vibrio cholerae* contains a divergent *recA* and provides multipathway protection from DNA damage. Environ Microbiol. 2015: 17(4):1090–1102. https://doi.org/10.1111/1462-2920.12512.

Rivera IN, Chun J, Huq A, Sack RB, Colwell RR. Genotypes associated with virulence in environmental isolates of *Vibrio cholerae*. Appl Environ Microbiol. 2001:67(6):2421–2429. https://doi.org/10.1128/AEM.67.6.2421-2429.2001.

Robins WP, Mekalanos JJ. Genomic science in understanding cholera outbreaks and evolution of *Vibrio cholerae* as a human pathogen. Curr Top Microbiol Immunol. 2014:379:211–229. https://doi.org/10.1007/82_2014_366.

Rousset F, Depardieu F, Miele S, Dowding J, Laval AL, Lieberman E, Garry D, Rocha EPC, Bernheim A, Bikard D. Phages and their satellites encode hotspots of antiviral systems. Cell Host Microbe. 2022:30(5):740–753.e5. https://doi.org/10.1016/j.chom.2022.02.018.

Roux S, Hallam SJ, Woyke T, Sullivan MB. Viral dark matter and virus-host interactions resolved from publicly available microbial genomes. Elife. 2015:4:e08490. https://doi.org/10.7554/eLife.08490.

Santoriello FJ, Michel L, Unterweger D, Pukatzki S. Pandemic *Vibrio cholerae* shuts down site-specific recombination to retain an interbacterial defence mechanism. Nat Commun. 2020:11(1):6246. https://doi.org/10.1038/s41467-020-20012-7.

Shapiro BJ, Polz MF. Ordering microbial diversity into ecologically and genetically cohesive units. Trends Microbiol. 2014:22(5):235–247. https://doi.org/10.1016/j.tim.2014.02.006.

Stutzmann S, Blokesch M. Circulation of a quorum-sensing-impaired variant of *Vibrio cholerae* strain C6706 masks important phenotypes. mSphere. 2016:1(3):e00098-16. https://doi.org/10.1128/mSphere.00098-16.

Stutzmann S, Blokesch M. Comparison of chitin-induced natural transformation in pandemic *Vibrio cholerae* O1 El Tor strains. Environ Microbiol. 2020:22(10):4149–4166. https://doi.org/10.1111/1462-2920.15214.

Suttle CA. Marine viruses—major players in the global ecosystem. Nat Rev Microbiol. 2007:5(10):801–812. https://doi.org/10.1038/nrmicro1750.

Tatusova T, DiCuccio M, Badretdin A, Chetvernin V, Nawrocki EP, Zaslavsky L, Lomsadze A, Pruitt KD, Borodovsky M, Ostell J. NCBI prokaryotic genome annotation pipeline. Nucleic Acids Res. 2016:44(14):6614–6624. https://doi.org/10.1093/nar/gkw569.

Taylor RK, Miller VL, Furlong DB, Mekalanos JJ. Use of *phoA* gene fusions to identify a pilus colonization factor coordinately regulated with cholera toxin. Proc Natl Acad Sci U S A. 1987:84(9): 2833–2837. https://doi.org/10.1073/pnas.84.9.2833.

Tesson F, Hervé A, Mordret E, Touchon M, d'Humieres C, Cury J, Bernheim A. Systematic and quantitative view of the antiviral arsenal of prokaryotes. Nat Commun. 2022:13(1):2561. https://doi.org/10.1038/s41467-022-30269-9.

Toussaint A, Chandler M. Prokaryote genome fluidity: toward a system approach of the mobilome. Methods Mol Biol. 2012:804:57–80. https://doi.org/10.1007/978-1-61779-361-5_4.

Tria FDK, Landan G, Dagan T. Phylogenetic rooting using minimal ancestor deviation. Nat Ecol Evol. 2017:1(7):193. https://doi.org/10.1038/s41559-017-0193.

van Houte S, Buckling A, Westra ER. Evolutionary ecology of prokaryotic immune mechanisms. Microbiol Mol Biol Rev. 2016:80(3): 745–763. https://doi.org/10.1128/MMBR.00011-16.

Vassallo CN, Doering CR, Littlehale ML, Teodoro GIC, Laub MT. A functional selection reveals previously undetected anti-phage defence systems in the *E. coli* pangenome. Nat Microbiol. 2022:7(10): 1568–1579. https://doi.org/10.1038/s41564-022-01219-4.

Vizzarro G, Lemopoulos A, Adams DW, Blokesch M. *Vibrio cholerae* pathogenicity island 2 encodes two distinct types of restriction systems. J Bacteriol. 2024:206(9):e0014524. https://doi.org/10.1128/jb.00145-24.

Wagner S, Grin I, Malmsheimer S, Singh N, Torres-Vargas CE, Westerhausen S. Bacterial type III secretion systems: a complex device for the delivery of bacterial effector proteins into eukaryotic host cells. FEMS Microbiol Lett. 2018:365(19):fny201. https://doi.org/10.1093/femsle/fny201.

Waldor MK, Mekalanos JJ. Lysogenic conversion by a filamentous phage encoding cholera toxin. Science. 1996:272(5270): 1910–1914. https://doi.org/10.1126/science.272.5270.1910.

Williams KP. Integration sites for genetic elements in prokaryotic tRNA and tmRNA genes: sublocation preference of integrase subfamilies. Nucleic Acids Res. 2002:30(4):866–875. https://doi.org/10.1093/nar/30.4.866.

Williams KP. Traffic at the tmRNA gene. J Bacteriol. 2003:185(3): 1059–1070. https://doi.org/10.1128/JB.185.3.1059-1070.2003.

Wozniak RA, Fouts DE, Spagnoletti M, Colombo MM, Ceccarelli D, Garriss G, Déry C, Burrus V, Waldor MK. Comparative ICE genomics: insights into the evolution of the SXT/R391 family of ICEs. PLoS Genet. 2009:5(12):e1000786. https://doi.org/10.1371/journal.pgen.1000786.

Yildiz FH, Schoolnik GK. Role of *rpoS* in stress survival and virulence of *Vibrio cholerae*. J Bacteriol. 1998:180(4):773–784. https://doi.org/10.1128/JB.180.4.773-784.1998.

Associate editor: Tal Dagan