Molecular insights into the genome dynamics and interactions between core and acquired genomes of *Vibrio cholerae*

Archana Pant[a,1], Satyabrata Bag[a,1,2], Bipasa Saha[a,1], Jyoti Verma[a,1], Pawan Kumar[a], Sayantan Banerjee[b], Bhoj Kumar[c], Yashwant Kumar[d], Anbumani Desigamani[a], Suhrid Maiti[e], Tushar K. Maiti[c], Sanjay K. Banerjee[d,f], Rupak K. Bhadra[g], Hemanta Koley[e], Shanta Dutta[e], G. Balakrish Nair[h,3], Thandavarayan Ramamurthy[a,e], and Bhabatosh Das[a,3]

[a]Molecular Genetics Laboratory, Infection and Immunology Division, Translational Health Science and Technology Institute, Faridabad 121001, India; [b]South-East Asia Regional Office, World Health Organization, New Delhi 110002, India; [c]Regional Centre for Biotechnology, National Capital Region Biotech Science Cluster, Faridabad 121001, India; [d]Non-communicable Diseases Division, Translational Health Science and Technology Institute, Faridabad 121001, India; [e]Scheme XM, National Institute of Cholera and Enteric Diseases, Beliaghata, Kolkata 700 010, India; [f]Department of Biotechnology, National Institute of Pharmaceutical Education and Research (NIPER-Guwahati), Changsari, Guwahati, Assam 781101, India; [g]Infectious Diseases and Immunology Division, Council of Scientific and Industrial Research Indian Institute of Chemical Biology, Kolkata 700032, India; and [h]Rajiv Gandhi Centre for Biotechnology, Thiruvananthapuram, Kerala 695014, India

Contributed by G. Balakrish Nair, June 26, 2020 (sent for review April 13, 2020; reviewed by Victor J. DiRita and Christophe Possoz)

Bacterial species are hosts to horizontally acquired mobile genetic elements (MGEs), which encode virulence, toxin, antimicrobial resistance, and other metabolic functions. The bipartite genome of *Vibrio cholerae* harbors sporadic and conserved MGEs that contribute in the disease development and survival of the pathogens. For a comprehensive understanding of dynamics of MGEs in the bacterial genome, we engineered the genome of *V. cholerae* and examined in vitro and in vivo stability of genomic islands (GIs), integrative conjugative elements (ICEs), and prophages. Recombinant vectors carrying the integration module of these GIs, ICE and CTXΦ, helped us to understand the efficiency of integrations of MGEs in the *V. cholerae* chromosome. We have deleted more than 250 acquired genes from 6 different loci in the *V. cholerae* chromosome and showed contribution of CTX prophage in the essentiality of SOS response master regulator LexA, which is otherwise not essential for viability in other bacteria, including *Escherichia coli*. In addition, we observed that the core genome-encoded RecA helps CTXΦ to bypass *V. cholerae* immunity and allow it to replicate in the host bacterium in the presence of similar prophage in the chromosome. Finally, our proteomics analysis reveals the importance of MGEs in modulating the levels of cellular proteome. This study engineered the genome of *V. cholerae* to remove all of the GIs, ICEs, and prophages and revealed important interactions between core and acquired genomes.

genomic islands | mobile genetic elements | site-specific integration | proteome | cholera

Bacterial genomes are highly heterogeneous, ranging from 500 to 10,000 kb in size, composed of single or multiple circular or linear chromosomes (Chr) with or without extra chromosomal genetic elements (1). The size and content of the bacterial genome are widely influenced by its lifestyle and living environments. Comprehensive analysis of genome sequences has shown the presence of number of horizontally acquired mobile genetic elements (MGEs) in the genomes of free living and facultative pathogens that contribute to virulence (2), colonization (3), invasions (4), antimicrobial resistance (5), xenobiotic degradation (6), and metabolic processes (7). Genomic islands (GIs), the most significant MGEs in the bacterial genome, are typically 10- to 200-kb discrete DNA segments that have distinct base composition compared to the whole genome, are distributed sporadically, encode mobility attributes, and are often linked with tRNA genes and flanked by direct repeat sequences (8). GIs are widely distributed in the genomes of both pathogenic and nonpathogenic bacterial species, are crucial for genome plasticity, and play important roles in the adaptation and evolution by encoding functions that do not exist in the core genome (9). GIs integrate to the host chromosome by site-specific recombination, can excise to form circular intermediates, and move into the genome of other bacteria via horizontal gene transfer (HGT) (10, 11). HGT is not restricted within the bacterial kingdom; it has become evident that MGEs present in the bacterial genomes can also move to the plant and animal genomes via lateral gene transfer (12).

Vibrio cholerae, the Gram-negative bacterial pathogen is responsible for the severe acute diarrheal disease cholera, harbors two nonhomologous circular chromosomes, chromosome 1 (Chr1) and chromosome 2 (Chr2) (13). Both the chromosomes are fortified with multiple prophages and GIs (Fig. 1). Two key virulence factors of *V. cholerae*, a pilus colonization factor toxin-coregulated pilus (TCP) and a potent enterotoxin cholera toxin (CT), which

Significance

Genome fluidity is directly associated with the emergence and evolution of many bacterial pathogens. Horizontally acquired genetic elements, which harbor fitness traits, are open source for the bacterial evolution. The genome of *Vibrio cholerae* is equipped with multiple mobile genetic elements (MGEs), which are crucial for disease development and survival. The present study engineered the genome of a well characterized clinical *V. cholerae* strain N16961 and reported (*i*) stability of MGEs in cholera pathogen, (*ii*) interactions between the core genome and MGEs, and (*iii*) role of MGEs in the modulation of cellular proteome. The engineered strain could be a potential candidate for understanding evolution of cholera pathogen and development of new therapeutic interventions.

Author contributions: G.B.N. and B.D. designed research; A.P., S. Bag, B.S., J.V., P.K., B.K., Y.K., A.D., S.M., and B.D. performed research; S. Banerjee, T.K.M., S.K.B., R.K.B., H.K., S.D., G.B.N., T.R., and B.D. contributed new reagents/analytic tools; A.P., S. Bag, B.S., J.V., S. Banerjee, G.B.N., T.R., and B.D. analyzed data; and G.B.N. and B.D. wrote the paper.

Reviewers: V.J.D., Michigan State University; C.P., CNRS.

The authors declare no competing interest.

[1]A.P., S. Bag, B.S., and J.V. contributed equally to this work.

[2]Present address: Department of Antimicrobial Resistance, National Centre for Disease Control, Delhi 110054, India.

[3]To whom correspondence may be addressed. Email: gbnair@rgcb.res.in or bhabatosh@thsti.res.in.

This article contains supporting information online at https://www.pnas.org/lookup/suppl/doi:10.1073/pnas.2006283117/-/DCSupplemental.

First published September 1, 2020.

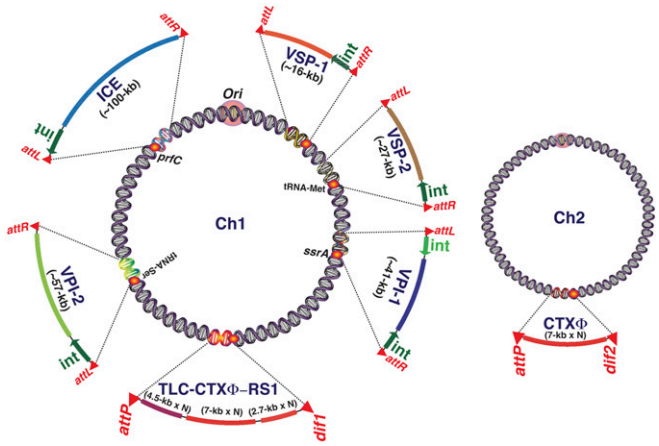

Fig. 1. Distribution of genomic islands (GIs), prophages, and integrative conjugative element (ICE) in the genome of *V. cholerae* clinical isolates. Each of the genetic elements except prophages is linked with different mobility functions like integrase and repeat sequences. GC contents of the acquired genetic elements are also distinct compared to the core genome. VPI-1, *Vibrio* pathogenicity island-1; VPI-2, *Vibrio* pathogenicity island-2; VSP-1, *Vibrio* seventh pandemic island-1; VSP-2, *Vibrio* seventh pandemic island-2; int, integrase.

helps the pathogen to colonize host intestinal epithelial cell surface and manifest the disease cholera, are part of the horizontally acquired *Vibrio* pathogenicity island-1 (VPI-1) and CTXΦ genome, respectively (2, 3). Genomic variability and clonal divergence of clinical and environmental *V. cholerae* isolates are mainly determined by the GIs (14–16). Following entry into host cytoplasm, GIs integrate into a specific site of chromosome using DNA recombinases (17, 18). The chromosome dimer resolution sites (*dif1* and *dif2*) of *V. cholerae* harbored several prophages known as integrative mobile elements exploiting Xer recombination (IMEX). CTXΦ and other vibriophages like RS1Φ, TLCΦ, and VGJΦ are well-known IMEXs (18). Of the several GIs, the VPI-1 is ubiquitous in the genome of all of the epidemic-causing *V. cholerae* isolates. VPI-1 integrates site-specifically into the transfer-messenger (tm) RNA (*ssrA*) locus of Chr1 (19). *Vibrio* pathogenicity island-2 (VPI-2), the GI that encodes functions for scavenging (*nanH*), transport (*dctPQM*), and catabolism (*nanA*, *nanE*, *nanK*, and *nagA*) of sialic acid, integrates into the tRNA-serine locus of Chr1 (20). The other GI, *Vibrio* seventh pandemic island-1 (VSP-1), encodes dinucleotide cyclase, the enzyme essential for producing intracellular signaling molecule cAMP-GMP, and integrates into the site between VC0174 and VC0186 in the strain N16961 (20). *Vibrio* seventh pandemic island-2 (VSP-2) is comparatively less stable and encodes RNase H1, DNA repair protein, methyl-accepting chemotaxis proteins, and type IV pilus and integrates into tRNA-methionine locus (13). Both the islands, VSP-1 and VSP-2, are consistently found in the genomes of seventh pandemic *V. cholerae* O1 El Tor and O139 strains, but are majorly absent in the genome of the first six pandemic strains recorded between 1817 and 1923 (21). SXT, a self-transmissible integrative conjugative element (ICE) that carries multiple antibiotic-resistance genes and confers resistance to sulfamethoxazole, trimethoprim, and streptomycin, was initially identified at the *prfC* locus in the genome of O139 *V. cholerae* (22). Among prophages integrated at the *dif* loci, CTX prophage is ubiquitously present in all of the epidemic strains. The toxin-linked cryptic phage (TLCΦ) and satellite phage RS1 are sporadic even in the clinical isolates (23). Except CTXΦ and RS1Φ, integrations of all of the GIs and SXT present in *V. cholerae* are reversible (20, 24). Although the excision frequency of GIs and SXT in *V. cholerae* is known, no study has been conducted to determine the frequency of loss of GIs under in vitro (test tube experiment) and in vivo (animal model) conditions. No effort so far has been made to construct derivatives of a clinical *V. cholerae* strain devoid of GIs, SXT, and/or prophages to understand the importance of each of these genetic elements in the adaptation and evolution of the cholera pathogen. More importantly, no study has been conducted to understand the importance of GIs in the essentiality of gene functions encoded by the core genome. Until now, we also have limited knowledge regarding how GIs in *V. cholerae* modulate the level of different cellular proteins.

In this study, we have engineered the genome of *V. cholerae* reference strain N16961 and its derivatives to monitor in vitro and in vivo stability of the GIs. We selectively or simultaneously deleted all of the GIs and prophages from the genome of N16961 and studied their integration and excision efficiency in the presence and absence of other GIs. We have also demonstrated how CTX prophage determines the essentiality of SOS master regulator LexA in *V. cholerae*, which is otherwise not essential in other bacterial species, including *Escherichia coli*. We have identified how chromosomally encoded function helps CTXΦ to bypass host immunity and replicate in the bacterial cell in the presence of similar prophages in the *V. cholerae* genome. Finally, we have analyzed whole-cell proteome of *V. cholerae* in the presence and absence of each of the GIs to unveil their contribution in modulating cellular metabolic processes. The knowledge generated from the present study will be of immense importance to understand the dynamics and importance of GIs in the adaptation and evolution of cholera pathogen and could be translated to develop *V. cholerae*-specific therapeutics to control cholera and other enteric diseases.

Results

Development of Reporter Strains to Monitor the Frequency of Loss of GIs in *V. cholerae*. Previous works that studied the excision phenomenon of GIs in *V. cholerae* adopted the PCR assay (19, 20). However, this assay is not proficient in monitoring the loss of GIs and unable to isolate bacterial cells from a heterogeneous population that have lost the GI from their genome. In this study, we engineered the genome of different *V. cholerae* strains to monitor the frequency of loss of different GIs under in vitro and in vivo growth conditions. Each GI (VPI-1, VPI-2, VSP-1, and VSP-2) and the prophages were tagged with multiple selectable (*cat*), counter-selectable (*sacB*), and chromogenic (*lacZ*) alleles in the clinical *V. cholerae* strain N16961 and its derivatives (Fig. 2). Since N16961 genome is devoid of SXT, we tagged this element with similar selectable and counter-selectable marker in a *V. cholerae* O1 El Tor clinical isolate 41081 (Table 1). The allele *sacB-cat* or *sacB-cat-lacZ* or *sacB-aadA* was introduced into the GIs or SXT by sequential allelic exchange method using derivatives of suicide vector pKAS32 harboring selectable marker *bla* and counter-selectable marker *rpsL* in the plasmid backbone (*SI Appendix*, Table S1). First cross-over intermediate strain carrying vector backbone (pKAS32) and reporter allele (*sacB-cat* or *sacB-cat-lacZ* or *sacB-aadA*) was confirmed by antibiotic resistance (Cam^R or Spc^R and Amp^R) and susceptibility to streptomycin and sucrose. The desired reporter strains SB30 (VPI-1::*sacB-cat*), SB26 (VPI-2::*sacB-cat*), SB23 (VSP-1::*sacB-cat*), DD1(VSP-2::*sacB-cat*), BB2 (CTXΦ::*sacB-cat*), and JV3 (SXT::*sacB-aadA*) were confirmed by antibiotic resistance, sucrose susceptibility, colony color, PCR assay, and DNA sequencing (Fig. 3). Since integration of CTXΦ and RS1Φ are irreversible, the TLC-CTX-RS1 prophage array in the genome of N16961 was first replaced by spectinomycin resistance cassette by allelic exchange method. Subsequently, CTXΦ::*sacB-cat* element was introduced in the *lacZ-dif1* allele by site-specific recombination to develop the recombinant bacterial strain BB2 (Table 1). Antibiotic resistance phenotype, PCR assay, and blue-white screening confirmed the desired strain BB2.

MICROBIOLOGY

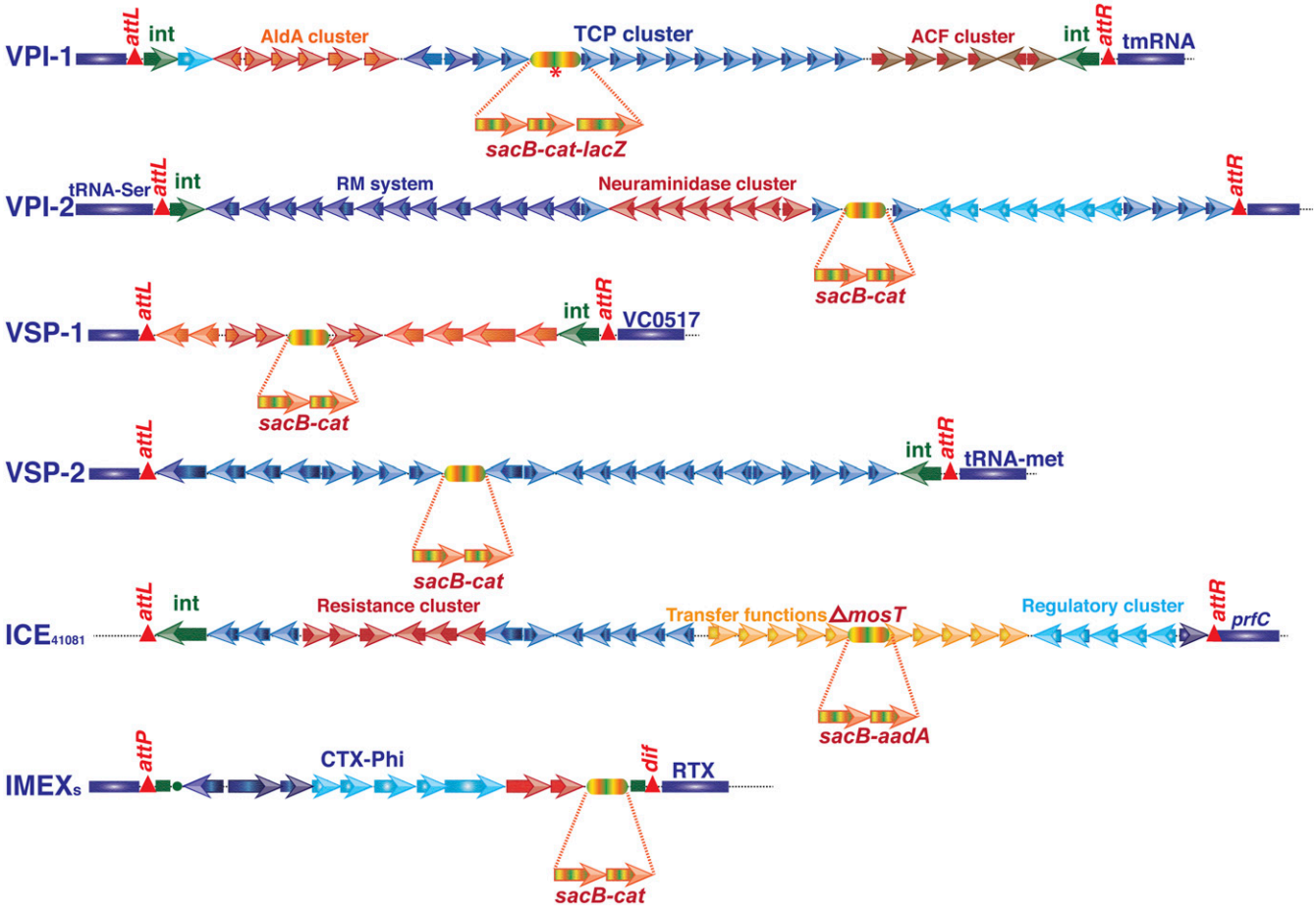

Fig. 2. Schematic representation of the genetic organization of the genomic islands (GIs), integrative conjugative element (ICE), and prophages and location of reporter genes in the tagged genome. Selectable (*cat*), counter-selectable (*sacB*), and chromogenic (*lacZ*) alleles were introduced in the *V. cholerae* genome, adopting an allelic exchange method. Asterisk indicates that *Vibrio* pathogenicity island-1 (VPI-1) was tagged with *sacB-cat* as well as *sacB-cat-lacZ* alleles.

Determination of In Vitro and In Vivo Frequency of Loss of GIs, Integrative Conjugative Element, and CTX Prophage. To monitor the stability of all of the four GIs, CTX prophage, and SXT element under in vitro growth conditions, the reporter strains engineered in this study (SB30, SB26, SB23, DD1, BB2, and JV3) were grown in nutritionally rich medium (Luria–Bertani broth; LB) for 12 h at 37 °C. Cultures of the reporter strains were plated onto Luria agar (LA) plates supplemented with sucrose (15%) and streptomycin (100 μg/mL). Colonies grown on the selection plates were monitored for their sensitivity to chloramphenicol or spectinomycin depending on the resistance cassette used to tag the GIs, CTX prophage, and SXT element. The excised colonies were further subjected to PCR confirmation by amplifying a chromosomal attachment region of each of the GIs (Fig. 3). Attachment sequences (*attB*) for each GI and SXT were confirmed by DNA sequencing. Total number of bacterial cells in an overnight-grown bacterial culture was determined by plating the same culture on an LA plate supplemented with streptomycin. We observed different frequency of loss for different GIs (Table 2). Highest frequency of loss under in vitro growth conditions was observed for the VPI-1, while no loss was detected for CTX prophage (Table 2).

To monitor the frequencies of loss of different GIs under in vivo growth conditions, we adopted the ligated rabbit ileal loop experimental model. Reporter strains harboring flagged GIs in their genome were inoculated into a ligated ileal loop, and the frequencies of loss of each GI were measured after 12 h by plating cells onto selection plates supplemented with sucrose (15%) and streptomycin (100 μg/mL). Like the in vitro experiment, total numbers of bacterial cells in the ligated ileal loop fluid were quantified by plating the fluid on the LA plate supplemented with streptomycin. The highest frequency of loss was observed for the VPI-2 compared to any other GIs (Table 2). No loss was observed for the CTX prophage (Table 2). Reporter strains carrying deletion of GIs isolated from ligated ileal loops were confirmed by restreaking the colonies onto the *Vibrio*-specific selective medium TCBS and LA supplemented with streptomycin (resistant) or chloramphenicol (sensitive). Interestingly, we observed that the frequency of loss of all of the GIs from the *V. cholerae* genome was very high under in vitro condition compared to in vivo-grown cells (Table 2).

***V. cholerae* Strains Devoid of Any Genomic Islands and Prophages Are Viable and Can Multiply both in Rich and Minimal Media.** We have done extensive genome engineering of N16961 to remove all its GIs and prophages from its genome. First, we tagged the GIs with *sacB-cat* or *sacB-cat-lacZ* and then isolated *V. cholerae* cells that lost the tagged GI by plating the cells on selection medium supplemented with sucrose (Table 1). To isolate strains lacking different combinations of MGE like ΔVPI-1 (SB31), ΔVPI-2 (SB27), ΔVSP-1 (SB25), ΔVSP-2 (DD2), ΔTLCΦ-CTXΦ-RS1Φ

Table 1. Relevant genotype and phenotype of wild-type and genetically modified *Vibrio cholerae* strains used in the study

Strain	Genotype and phenotype	Reference
N16961	Wild type, O1 El Tor, Stp^r	13
O395	Wild type, O1 classical, Stp^r	5
VCE232	Wild type, O4 environmental	5
41081	O1 El Tor, SXT+ve, NA^r, Fun^r, Plm^r	MVIDH, India
MO10	Wild type, O139 serogroup, Stp^r, SXT$^+$	22
BS1	N16961ΔTLCΦ-CTXΦ-*dif1*::*lacZ-dif1*Δ*dif2*, Stp^r, Spc^r	25
BS11	BS1Δ*recA*, Stp^r, Spc^r	25
AP7	N16961Δ*recA*, Stp^r	This study
AP8	VCE232Δ*recA*::*aph*, Kan^r	This study
AP10	O395Δ*recA*::*aph*, Stp^r, Kan^r	This study
SB23	N16961*vsp1*::*sacB-cat*, Stp^r, Cam^r, *Suc*^{s}	This study
SB25	N16961*Δvsp1*, Stp^r	This study
DD1	N16961*vsp2*::*sacB-cat*, Stp^r, Cam^r	This study
DD2	N16961*Δvsp2*, Stp^r	This study
SB30	N16961*vpi1*::*sacB-cat*, Stp^r, Cam^r, *Suc*^{s}	This study
SB31	N16961*Δvpi1*, Stp^r	This study
SB26	N16961*vpi2*::*sacB-cat*, Stp^r, Cam^r, *Suc*^{s}	This study
SB27	N16961*Δvpi2*, Stp^r	This study
SB32	N16961*Δvsp1 vsp2*::*sacB-cat*, Stp^r, Cam^r, *Suc*^{s}	This study
SB33	N16961*Δvsp1 Δvsp2*, Stp^r	This study
SB42	N16961*Δvpi2*, *ΔlacZ*, Stp^r	This study
SB43	N16961*Δvpi1*, *ΔlacZ*, Stp^r	This study
SB44	N16961*Δvsp1 Δvsp2 ΔlacZ*, Stp^r	This study
SB51	N16961*Δvpi2ΔlacZvpi-1*::*sacB-cat-lacZ*, Stp^r, Cam^r, *Suc*^{s}	This study
SB52	N16961 *Δvpi-1 Δvpi-2 ΔlacZ*, Stp^r	This study
SB45	N16961*Δvsp1Δvsp2ΔlacZvpi1*::*sacB-cat-lacZ*, Stp^r, Cam^r, *Suc*^{s}	This study
SB47	N16961*Δvsp1Δvsp2 Δvpi1 ΔlacZ*, Stp^r	This study
SB46	N16961*Δvsp1Δvsp2ΔlacZvpi-2*::*sacB-cat-lacZ*, Stp^r, Cam^r, *Suc*^{s}	This study
SB48	N16961 *Δvsp1 Δvsp2Δ vpi-2 ΔlacZ*, Stp^r	This study
SB49	N16961*Δvsp1Δvsp2Δvpi1ΔlacZvpi2*::*sacB-cat-lacZ*, Stp^r, Cam^r, *Suc*^{s}	This study
SB50	N16961*Δvsp1Δvsp2 Δvpi1Δvpi2 ΔlacZ*, Stp^r	This study
SB53	N16961Δ*vpi1*Δ*vpi2*Δ*vsp1*Δ*vsp2*Δ*lacZ*ΔTLCΦ-RS1-CTXΦ-*dif1*::*aadA*, Stp^r, Spec^r	This study
SB58	N16961*Δvsp1*::*pBD60*, Stp^r, Cam^r	This study
SB59	N16961*Δvsp2*:: *pBD60*, Stp^r, Cam^r	This study
SB56	N16961::*cat*, Stp^r, Cam^r	This study
SB60	N16961 *Δvsp1 Δvsp2*::*cat*, Stp^r, Cam^r	This study
SB62	O395::*sh ble*, Stp^r, Zeo^r	This study
SB63	N16961 *Δvpi-1 Δvpi-2*::*pBD60*, Stp^r, Cam^r	This study
BB1	N16961Δ*vpi1*Δ*vpi2*Δ*vsp1*Δ*vsp2*Δ*lacZ*ΔTLCΦ-RS1-CTXΦ-*dif1*::*lacZ-dif1*	This study
BB2	N16961Δ*vpi1*Δ*vpi2*Δ*vsp1*Δ*vsp2*Δ*lacZ*ΔTLCΦ-RS1-CTXΦ-*dif1*::*lacZ-dif1-CTXΦ-sacB-cat*	This study
BS20	BS1Δ*lexA*, Stp^r, Spc^r, Kan^r	26
BD14	BS20 *lexA(Ind $^-$)*	26
JV2	41081SXT:Δ*strAB*, Stp^r NA^r, Fun^r, Plm^r	This study
JV3	41081SXT:Δ*strAB*::*sacB-aadA*, Stp^r NA^r, Fun^r, Plm^r, Spec^r	This study
JV4	41081ΔSXT, Stp^r, NA^r, Fun^r, Plm^r	This study
JV5	BS1Δ*lexA*Δ*recA*, Stp^r, Spc^r, Kan^r	This study

(BS1), ΔVPI-1ΔVPI-2 (SB52), ΔVSP-1ΔVSP-2 (SB44), ΔVSP-1ΔVSP-2ΔVPI-1 (SB47), ΔVSP-1ΔVSP-2ΔVPI-2 (SB48), ΔVSP-1ΔVSP-2ΔVPI-1ΔVPI-2 (SB50), and ΔVSP-1ΔVSP-2ΔVPI-1ΔVPI-2ΔTLCΦ-CTXΦ-RS1Φ::*lacZ-dif1* (BB1) strain (Table 1), we utilized conditional toxicity of *sacB* to the host bacteria.

We measured the growth rate of N16961 and its derivatives with or without GIs and prophages in a nutritionally rich (LB) or minimal (M9) medium (nutritionally poor). *V. cholerae* cells devoid of any GIs and prophages were able to grow in both media (Fig. 4 *A* and *B*). However, growth of SB50 (Δ*vsp-1*Δ*vsp-2*Δ*vpi-1*Δ*vpi-2*) was significantly decreased ($P < 0.01$) during 5, 6, and 7 h of growth when compared with its parent strain N16961 (Fig. 4*A*). We also observed significant reduction ($P < 0.05$) in growth of the engineered strain SB47 (Δ*vsp-1*Δ*vsp-2* Δ*vpi-1*) during 5, 6, and 7 h of incubation. Interestingly, BB1, the derivative of SB50 carrying *dif1* in the Chr1 but devoid of all of the four GIs and prophages in the chromosome, showed a similar growth profile as the WT strain N16961 (Fig. 4 *A* and *B*). We have not observed any major changes in growth profile either in M9 or LB between N16961 and BB1 strains (Fig. 4 *A* and *B*). It has previously been reported that the *V. cholerae* O1 El Tor strain N16961 has a significant growth advantage over the classical biotype strain O395 in stationary-phase culture when both strains were cocultured in LB under optimal growth conditions (27). The major difference between El Tor and classical biotypes with respect to GI is the absence of both VSP-1 and VSP-2 in the latter. We investigated whether O1 classical strain O395 can compete with O1 El Tor strain SB60 devoid of VSP-1 and VSP-2

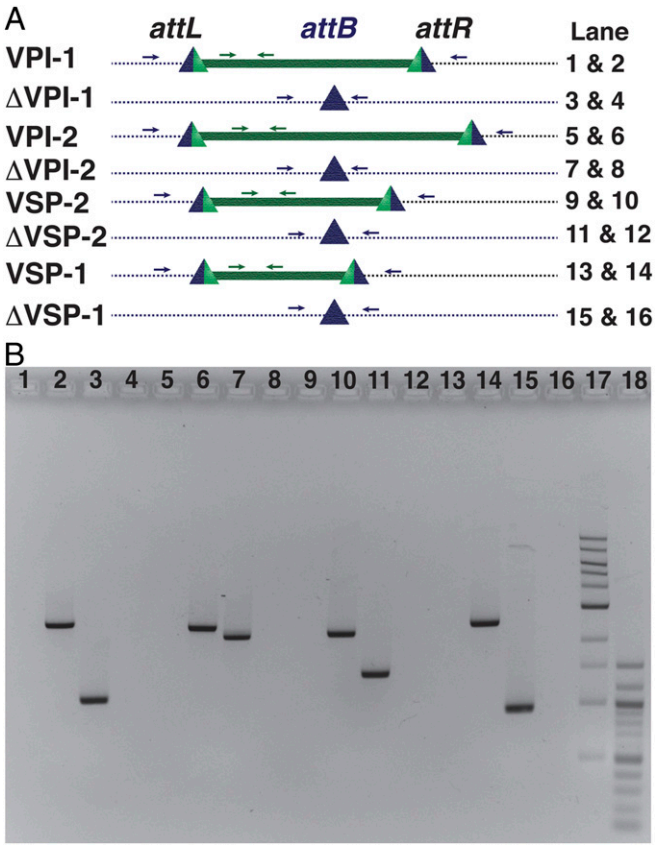

Fig. 3. Genetic organization of chromosomal integration sites (*attB*) before and after excision of all of four genomic islands from the genome of N16961. (*A*) *attL* and *attR* sites were formed due to site-specific recombination between genomic island attachment site (*attP*) and chromosomal attachment site (*attB*). Lane numbers 17 and 18 contained 1-kb and 100-bp DNA ladders, respectively. (*B*) Polymerase chain reaction (PCR) analysis of N16961 and its derivatives in the presence and absence of genomic islands (GIs). Amplified PCR products were resolved in an agarose gel. Lane number and corresponding amplicon are mentioned in *A*. Blue arrows indicate primers that bind to the upstream and downstream of GI, and green arrows indicate primers that bind to the GI.

islands. To differentiate the SB60 and O395 strains in the coculture, we introduced two different resistance gene cassettes, *cat* and *sh ble*, in the *dif1* site of El Tor and classical strains, respectively. The mixed culture was grown in LB or M9 medium under optimal growth conditions, and the number of bacterial cells was measured at different time points (Fig. 4 *C* and *D*). Our finding indicates that O395 classical biotype can compete with the VSP-1 and VSP-2–deleted El Tor biotype when they were cocultured in M9 medium. However, the VSP-1– and VSP-2–deleted El Tor strains can outcompete O395 in the LB medium. This finding indicates that the acquisition of VSP-1 and VSP-2 in the genome of *V. cholerae* provided certain metabolic traits that possibly helped the seventh pandemic El Tor biotype to outcompete the sixth pandemic-causing classical biotype. However, additional role of *oriC* functions might also help the El biotype to outcompete the classical biotype in nutrient-rich environment even in the absence of such metabolic traits (28). It has previously been reported that the *oriC* of the classical biotype carries a single base mutation (T→G) in its conserved AT-rich 13-mer R repeat region compared to the *oriC* of El Tor biotype (28). It has been demonstrated that the minichromosome copy number and its transformational efficiency are substantially reduced due to substitution mutation in the *oriC* of the classical biotype (28). This might be one of the reasons for less competency of classical biotype while coculturing with the El Tor biotype in the nutrient-rich medium.

Integration Efficiency of Different GIs, SXT, and CTXΦ in the Presence and Absence of Acquired Genome. Studies on integration efficiency of GIs are complicated because of their large size and nonreplicative nature of their excised circular form. Therefore, to avoid this complexity, we amplified the integrative module of GIs, SXT, and CTXΦ from its excised circular form, which includes attachment site (*attP*) with or without recombinases (Fig. 5). The integration modules were cloned in a conditionally replicative conjugative vector (*SI Appendix*, Table S1). Recombinant vector containing integration module of GIs, SXT, or CTXΦ was introduced by conjugation into *V. cholerae* strains carrying native chromosomal attachment site (*attB*) with or without other GIs. Integration of integrative module in the chromosome of *V. cholerae* was determined based on the antibiotic (chloramphenicol or zeocin) resistance profile of recipient cells. Integration of a GI at specific chromosomal locus was further confirmed by PCR assay and blue-white colorimetric assay. Conjugation of different recombinant vectors carrying nonreplicative integration module of different GIs, SXT, and CTXΦ gave rise to different numbers of transconjugants (Table 3). Among nonreplicative GIs and SXT, the highest number of transconjugants was obtained with the recombinant vector carrying the integration module of SXT (Table 3). Absence of GIs and prophages increased the integration efficiency of SXT. A minimal number of transconjugants was obtained with recombinant vector carrying the integration module of VSP-1 (Table 3). However, when compared with the integration efficiency of GIs and SXT with the replicative and integrative module of CTXΦ, the highest number of transconjugants was obtained with the phage genome (Table 3). To differentiate between *V. cholerae* cells carrying the integrative or replicative form of the RS2 module of CTXΦ, we fused the chromosomal attachment site *dif1* in the coding region of the *lacZ* gene amplified from *E. coli*. The *lacZ-dif1* allele produced functional β-galactosidase and generated blue colonies in the presence of

Table 2. In vitro and in vivo loss frequency of different genomic islands

Genomic Islands/ICE/Prophages	Bacterial strain	Loss frequency, mean ± SEM	
		In vitro	In vivo*
VPI-1::*sacB-cat*	SB30	3.26×10^{-4} (±0.00007)	1.03×10^{-9}
VPI-2::*sacB-cat*	SB26	1.40×10^{-4} (±0.00005)	1.20×10^{-8}
VSP-1::*sacB-cat*	SB23	4.21×10^{-5} (±0.00001)	2.11×10^{-9}
VSP-2::*sacB-cat*	DD1	1.41×10^{-4} (±0.00001)	1.08×10^{-9}
SXT::*sacB-aadA*	JV3	1.53×10^{-4} (±0.00005)	ND
CTXϕ:: *sacB-cat*	BS37	$<10^{-10}$	$<10^{-10}$

*In vivo rabbit ileal loop experiment was done in duplicate set.

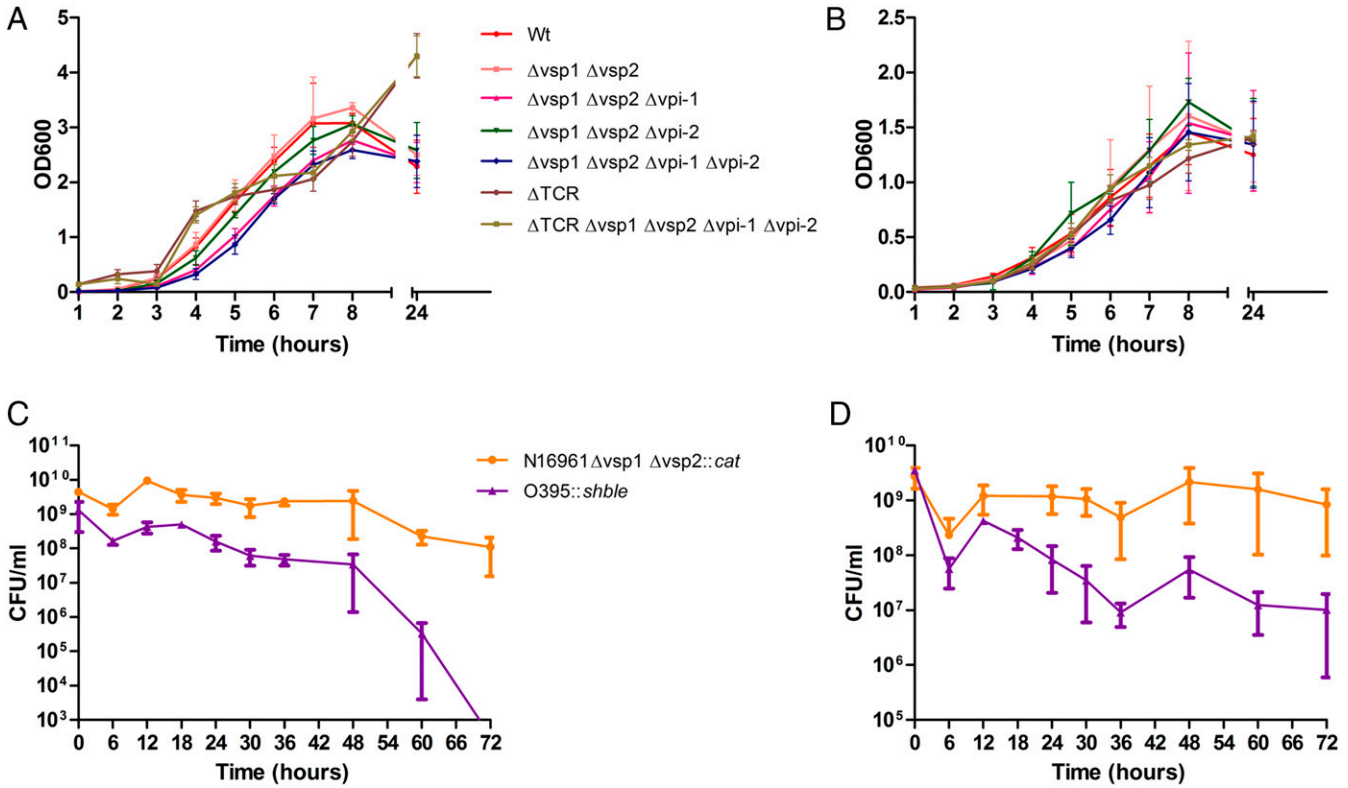

Fig. 4. Growth curve of wild type and genetically engineered *V. cholerae* strains in rich (LB) and nutritionally defined culture medium (M9). The El Tor strain N16961 and its derivatives were grown individually in monocultures in LB (*A*) and M9 (*B*), and the growth of the bacterial cells was monitored by measuring the optical density of the culture at 600 nm (OD_{600}) in a spectrophotometer at different time points. Growth competition between ΔVSP-1ΔVSP-2 N16961 and O395 was also done in rich (*C*) and minimal medium (*D*) in a cocultured medium by inoculating similar number of cells together (1:1). CFU count of each strain was determined by plating the culture in rich medium (LA) supplemented with specific antibiotic. TCR = TLCΦ-CTXΦ-RS1Φ.

chromogenic substance X-gal in the selection plate. We then deleted endogenous *lacZ* gene from *V. cholerae* cells, and the TLC-CTX-RS1 prophage array was replaced with *lacZ-dif1* allele (*SI Appendix*, Fig. S3). The reporter strain carrying the replicative RS2 module of CTXΦ turned blue in the presence of X-gal on the selection plate, while the strain carrying integrative RS2 in the *dif1* in the chromosomal *lacZ-dif1* allele produced white colonies due to inactivation of *lacZ* gene and abolishing β-galactosidase production (*SI Appendix*, Fig. S3). The number of transconjugants with same recombinant vector in the presence and absence of other GIs in the host cells showed different outcomes (Table 3). The integration efficiencies of CTXΦ and its derivatives were measured in the absence of TLCΦ and other prophages. The integration module of SXT and VPI-1 can integrate almost with equal efficiency in the presence and absence of other GIs in the *V. cholerae* genome. Integration efficiency of VPI-2 was reduced in the absence of other GIs. Interestingly, integration efficiency of VSP-1 and VSP-2 was increased in the absence of GIs (Table 3).

We have also measured the integration efficiency of the nonreplicative RS2 module of CTXΦ in the *dif1* site of Chr1 in *V. cholerae* reporter strain BS1 and BB1. We observed similar integration efficiency of nonreplicative RS2 module in the presence and absence of different GIs. The integration event of nonreplicative RS2 module was site-specific (≥99.3%) and irreversible, both in the presence and absence of other GIs (Table 3).

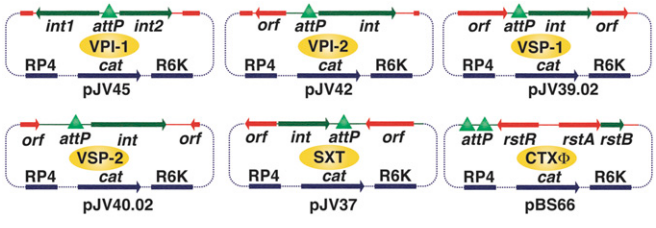

Fig. 5. Genetic tools engineered to measure the integration efficiency of different GIs, ICE (SXT), and CTXΦ. The integration module of all of the elements was amplified from their excised circular form, which include attachment site (*attP*) with or without tyrosine recombinases. The original vector (pSW23T) was unable to replicate in *V. cholerae* due to its conditional *ori* (R6K) function.

Cross-Talk between Horizontally Acquired Genetic Elements and Core Genome. Cross-talk between GIs and prophages in *V. cholerae* is a well-known phenomenon (29, 30). However, the influence of acquired functions in the functionality of genes present in the core genome is not known. Understanding the molecular basis of essentiality of LexA in *V. cholerae* is of special interest because of its contribution in the biology of horizontally acquired MGEs, especially in the replication, integration, and production of extrachromosomal CTXΦ from prophages and excision and dissemination of SXT. We resolved the mystery of LexA essentiality in *V. cholerae* by deleting the *lexA* gene from the chromosome in the engineered *V. cholerae* strain BB1 deleted with GIs and prophages (Table 1). We utilized a recombinant vector pAP17 containing kanamycin resistance cassette (*aph3*) flanked by the upstream and downstream regions of *lexA* to delete the *lexA* gene by allelic exchange method. To identify the key genetic element

Table 3. Integration frequencies of different genomic islands, CTXΦ and SXT element in the presence and absence of acquired genetic elements

Genetic element	Bacterial strain	Integration frequency (SD)
SXT	N16961	1.56×10^{-1} (0.0814)
	BB1	3.13×10^{-1} (0.1808)
VPI-1	SB31	1.3×10^{-3} (0.0005)
	BB1	1.3×10^{-3} (0.0005)
VPI-2	SB27	3.6×10^{-3} (0.0020)
	BB1	2.6×10^{-4} (0.0002)
VSP-I	SB25	9.3×10^{-6} (1.1547E-06)
	BB1	5×10^{-5} (1.90351E-05)
VSP-II	DD2	8.9×10^{-3} (0.0077)
	BB1	1.08×10^{-2} (0.0060)
CTXΦ	BS1	2.2×10^{-3} (0.0012)
	BB1	2.9×10^{-2} (0.046)

responsible for essentiality of the *lexA* gene in *V. cholerae* genome, we selected single or multiple GI or prophage-deleted strains (Table 1) followed by an attempt to delete *lexA* in each of these engineered strains. Surprisingly, such screening allowed us to discover that the *lexA* gene is not essential in TLC -CTX-RS1 prophage-deleted *V. cholerae* cells, but it is still essential in single or multiple GI-negative strain (Table 4). To determine whether TLC-CTX-RS1 array or any particular one of these phages is responsible for the essentiality of *lexA*, each of them was sequentially deleted, followed by an attempt to delete *lexA* in those engineered strains. Interestingly, it was identified that *lexA* is not essential only in CTX prophage-deleted cells (Table 4). CTX prophage is composed of RS2 and core region. While RS2 carries *rstR-rstA-rstB* genes, core segment carries seven different genes including CT-coding *ctxAB* genes. On the other hand, the RS1Φ is identical to the RS2 except it carries an extra gene, called *rstC*. The *rstA*, *rstB*, and *rstR* gene products are essential for phage replication, integration, and transcription regulation, respectively. Since the expression of *rstR*, *rstA*, and *rstB* is under the control of LexA from P_{rstA} promoter (31), we therefore analyzed the function of each of these proteins to understand their contribution, if any, in the essentiality of *lexA* gene in *V. cholerae*. RstB is a single-stranded DNA binding protein, which helps in integration of CTXΦ and the RS1Φ at the chromosomal *dif* site. RstR is a transcriptional regulator, which represses the transcription of *rstA*, *rstB*, and *rstC* from the P_{rstA} promoter (*SI Appendix*, Fig. S3). RstA is a Rep protein, which introduces DNA nicks in the ori_{ctx} sequence and helps CTXΦ to generate extrachromosomal-phage genome from the prophage. We hypothesized that, in the absence of LexA-mediated repression, there could be excess production of RstA from the CTX and RS1 prophages, which in turn generates frequent nicks in the CTX and RS1 prophage DNA, leading to lethal chromosomal breaks. To test this hypothesis, we tried to integrate CTXΦ or CTXΦ-RS1Φ in tandem in *V. cholerae* Δ*lexA* ΔCTXΔRS1 mutant but failed to integrate either multiple copies of CTXΦs or CTXΦ-RS1Φ in tandem. In sharp contrast, integration of multiple copies of CTXΦs or CTXΦ-RS1Φ in tandem in the *dif1* site of the *V. cholerae* ΔCTXΔRS1 but *lexA*-positive strain was easily obtained (*SI Appendix*, Fig. S1). Since RecA is intricately associated with LexA stability and thus prophage gene expression, we engineered a Δ*lexA*Δ*recA* double-mutant strain JV5 and investigated the CTX prophage number in the *dif1* site after introducing the phage by conjugation. Like Δ*lexA* strain, the single-copy CTX prophage was also dominant in Δ*lexA*Δ*recA* double mutant (*SI Appendix*, Fig. S1). To provide further evidence that the multiple copy of functional *rstA* is associated with the essentiality of LexA in *V. cholerae*, we performed site-directed mutagenesis to replace the catalytic tyrosine residue of the RIYNK motif of RstA protein with phenylalanine residue. Due to this point mutation, the engineered CTXΦ encoding catalytically inactive $RstA^{Y237F}$ mutant protein was unable to initiate rolling circle replication in *V. cholerae*. However, a single copy of CTXΦ carrying nonfunctional RstA was stably lysogenized both in the presence and absence of LexA. These results thus strongly suggest that, when functional *rstA* gene is present in *V. cholerae* genome, deletion of *lexA* is lethal to the bacterium. Monitoring of phenotypes (antibiotic resistance) as well as genotypes (PCR assay and sequencing) confirmed the authenticity of these *V. cholerae* mutants.

Core Genome-Encoded RecA Helps CTXΦ to Overcome Host Immunity and Reinfect *V. cholerae*. Other than phage-encoded RstA, chromosomally encoded UvrD helicase and LexA also play an important role in CTXΦ replication (30). It was reported that RstR, the CTXΦ encoded transcriptional factor, binds to the *rstA* promoter (P_{rstA}) and represses its expression from the prophage genome (31). However, complete repression occurs by cooperative interaction with LexA, the master regulator of bacterial SOS response. The importance of RecA, the second crucial protein for the bacterial SOS response, in the replication of CTXΦ in *V. cholerae* has not been explored. We measured the replication efficiency of El Tor, classical, and environmental types of CTXΦs in the presence and absence of RecA in different *V. cholerae* strains (Fig. 6). The genome of N16961 harbors a $CTX\Phi^{ET}$ prophage in its large chromosome. We introduced the replication and integration module of $CTX\Phi^{ET}$ ($RS2^{ET}$) and $CTX\Phi^{CI}$ ($RS2^{CI}$) into RecA-positive N16961 and isolated transconjugants on a selection plate supplemented with chloramphenicol. We obtained 2.44×10^4 and 3.32×10^5 transconjugants carrying $RS2^{ET}$ (pBS66) and $RS2^{CI}$ (pBS22), respectively (Fig. 6). We had also introduced complete $CTX\Phi^{ET}$, $RS2^{ET}$, $RS2^{CI}$, $RS2^{Env}$, and RS1Φ into a reporter *V. cholerae* strain BS1 lacking the CTXΦ prophage in its genome, which has

Table 4. Essentiality of *lexA* for the viability of *Vibrio cholerae* in the presence and absence of genomic islands and prophages

Strain	Genotype	Frequency of deletion of *lexA*	*lexA* essentiality
N16961	Wild type	$<1.0 \times 10^{-10}$	Yes
SB31	ΔVPI-1	$<1.14 \times 10^{-9}$	Yes
SB27	ΔVPI-2	$<1.03 \times 10^{-9}$	Yes
SB25	ΔVSP-1	$<2.91 \times 10^{-9}$	Yes
DD2	ΔVSP-2	$<2.1 \times 10^{-9}$	Yes
SB50	ΔVPI-1ΔVPI-2ΔVSP-1ΔVSP-2	$<4.0 \times 10^{-9}$	Yes
B33	ΔTLCΔRS1	$<3.92 \times 10^{-10}$	Yes
SB53	ΔVPI-1ΔVPI-2ΔVSP-1ΔVSP-2 ΔTLCΔCTXΔRS1	4.93×10^{-7}	No
BS1	ΔTLCΔCTXΔRS1	2.59×10^{-5}	No

Frequency of lexA deletion was measured from at least two independent experiments.

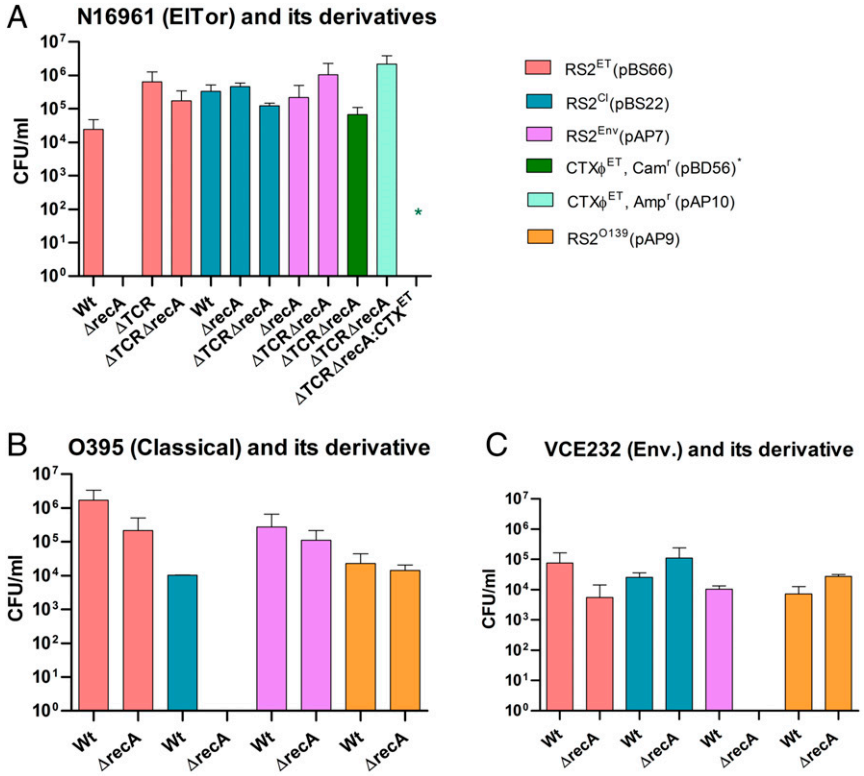

Fig. 6. Analysis of the replication efficiency of different CTXΦ in the presence and absence of RecA function. The bars represent numbers of CFUs carrying replicative module of CTXΦ. Replicative modules were transferred to the host bacteria by 3 h conjugation. Transconjugants were selected based on the antibiotic-resistance phenotypes. Two-way RM-ANOVA was used to infer statistical significance of the differences of CFU numbers. Error bars indicate SDs. Wt, wild type; ET, El Tor (*A*); Cl, classical (*B*); Env, environmental (*C*); TCR, TLC-CTX-RS1 prophage. * = CTXΦ^{ET}Cam.

a functional RecA. It is noteworthy to mention that similar numbers of transconjugants were obtained for each of the phages introduced (Fig. 6). When we introduced each of these phages in the BS1 derivative harboring CTXΦ^{ET} prophage at the *dif1* loci, here also we obtained a similar number of transconjugants carrying replicative and integrative phage genomes or RS modules. The results indicate that reinfection by CTXΦ is possible even in the presence of biotype-specific *rstR* and functional *recA* genes in the host genome (Fig. 6).

Since the function of RecA is intricately associated with the SOS response and modulates LexA autoproteolytic activity, we extended our study to decipher the role of RecA in CTXΦ replication. To do this, all of the genetic elements, namely CTXΦ^{ET}, RS2^{ET}, RS2^{Cl}, and RS2^{Env}, were introduced into the *recA*-deleted BS1 strain, and the numbers of transconjugants on the selection plates were compared. Similar numbers of transconjugants carrying replicative and integrative phage genomes or RS modules were obtained in the presence of functional RecA (*SI Appendix*, Table S2). We then deleted the *recA* gene in N16961 and introduced CTXΦ^{ET}, RS2^{ET}, RS2^{Cl}, and RS2^{Env} to monitor replication efficiency of different phage-derived elements. To our surprise, while no transconjugant was obtained for the CTXΦ^{ET} or RS2^{ET} element in the absence of *recA*, transconjugants for RS2^{Cl} (4.6×10^5) or RS2^{Env} (5×10^4) element were obtainable using the same *V. cholerae* mutant strain (Fig. 6). We further confirmed this phenotype using the N16961-derived BS11 strain, which lacks *recA* and all of the prophages. Once CTXΦ^{ET} was introduced, it also showed immunity against reinfection with the similar type of CTXΦ (Fig. 6). We also observed similar phenotypes in classical (O395) and environmental (VCE232) *V. cholerae* strains when we introduced CTXΦ^{Cl} or CTXΦ^{Env} replication and integration module to the strains lacking *recA* gene but containing CTXΦ^{Cl} or CTXΦ^{Env} in the chromosome, respectively (Fig. 6 and *SI Appendix*, Table S2). In both cases, CTXΦ can replicate in the absence of *recA* but presence of different biotypes of CTX prophages in the chromosome. We also ensured that the transmission process is not reduced in the absence of *recA* gene function by using a replicative conjugative plasmid pFX497 (*SI Appendix*, Table S1) as a control, which gave a similar number of transconjugants in the presence or absence of *recA*. However, RecA function was not linked with the integration efficiency of CTXΦ, since it is able to integrate at the *dif* site with almost equal efficiency in presence or absence of RecA function (*SI Appendix*, Table S2). The findings of the present study demonstrate that RecA helps CTXΦ replication in the presence of preexisting CTX prophage in the host chromosome.

Contribution of GIs in Modulating Cellular Proteome. For identification, quantification, and comparison of whole-cell proteomes of wild type *V. cholerae* strain N16961 and its derivative BB1 devoid of GIs and prophages, we performed experiments using a high-throughput liquid chromatography (LC) system (Eksigent 2D) coupled with Triple TOF 5600 mass spectrometer (AB Sciex). We have identified a total of 1,864 proteins in both strains (Dataset S1). Based on the proteome profile of N16961 and BB1, we categorized the protein pool into three classes: (*i*) proteins with lower abundance (>1.5 log fold changes) or absent in BB1, (*ii*) proteins with higher abundance (>1.5 log fold changes) in BB1, and (*iii*) proteins with similar abundance in both strains. Proteins that are not detectable in BB1 but found in the N16961 strain were analyzed to find their coding sequences. We have identified similar numbers of proteins in N16961

(n = 1,765) and BB1 (n = 1,724) strains. A total of 141 proteins were detected only in N16961, of which 41 proteins are encoded by the genes present in Chr2. More than 10 proteins are encoded by the 4 GIs, VPI-1 (n = 1), VPI-2 (n = 4), VSP-1 (n = 3), VSP-2 (n = 1), and CTXΦ (n = 2), which were detected in N16961 but not in BB1. We observed that 131 proteins were differentially expressed (log fold change 1.5 and P value <0.05) between N16961 and BB1 strains. Out of 131 differentially expressed proteins in BB1, 59 and 72 proteins were up- and down-regulated, respectively. Proteins down-regulated in BB1 belong to uncharacterized proteins (n = 8), outer membrane proteins (n = 2), flagellar biosynthetic proteins (n = 2), aspartate modifying enzymes (n = 2), and several others. Enzymes involved in succinate metabolic pathways (n = 3), RNA degradation (n = 2), and ABC transporters (n = 2) are up-regulated in the absence of MGEs. Both small- and large-subunit ribosomal proteins are up- and down-regulated in the absence of MGEs. We observed that 61 proteins encoded by the Chr2 were detectable either in N16961 (n = 25) or BB1 (n = 36).

Discussion

Acquisition of virulence functions among enteric pathogens through lateral gene transfer, loss of redundant and nonessential genes among intracellular pathogens, and antigenic drift among mucosal pathogens are well recognized evidences in the post-genomic era (32). A hallmark of evolution of the cholera pathogen is its high proficiency in genetic exchange and genomic plasticity. *V. cholerae* is naturally competent, and different pathways, including transformation, conjugation, and transduction can facilitate HGT in the pathogen. Comparative analysis of genome sequences of thousands of *V. cholerae* isolates collected over a period of 170 y revealed its genome plasticity and continuous evolution due to acquisition or loss of specific genomic segments, specifically bacteriophages, ICEs, GIs, and plasmids (14, 21, 33). MGEs carrying toxin, virulence factors, antimicrobial resistance function, metabolic enzymes, and other fitness factors are ubiquitously present in the genome of toxigenic *V. cholerae* strains. The genus *Vibrio* and other members of Vibrionaceae commonly harbor two nonhomologous circular chromosomes, Chr1 and Chr2 (34). However, a recent report indicates that *V. cholerae* might carry a single fused chromosome instead of two chromosomes (35). Exploration of the bipartite and single chromosome-containing genomes of clinical *V. cholerae* isolates identified four non–self-mobilizing genomic islands (VPI-1, VPI-2, VSP-1, VSP-2) and one self-mobilizing ICE and multiple prophages (CTXΦ, RS1Φ, TLCΦ) in most of the seventh pandemic isolates. CTXΦ, VPI-1, and VPI-2 are ubiquitously present in all of the epidemic isolates, whereas VSP-1 and VSP-2 islands are commonly present in the seventh pandemic O1 and O139 isolates (14, 33). Except IMEXs, all of the GIs and ICEs reported in *V. cholerae* encode their own tyrosine recombinases as predominant integrases and integrate reversibly in a specific chromosomal locus by site-specific recombination (19, 20). Changes in the bacterial genome sequences due to accumulation of spontaneous point mutations are of the order of 10^{-9} to 10^{-7} nucleotides per year (36). Analysis of multiple genomes of *V. cholerae* isolated from the same epidemic or even from the same patient found that genomic contents of the isolates may differ from less than 10 to more than 100 genes within each patient (37). Presence or absence of strain-specific DNA segments is found in the specific loci of either chromosome of *V. cholerae*.

Compared to acquisition, deletion events of genomic segments are highly variable in *V. cholerae* and other bacterial species, including *E. coli* (38). The genomic diversities among clinical and nontoxigenic environmental *V. cholerae* strains are also very high (39). It has been found that, like clinical isolates, the genomes of the environmental *V. cholerae* strains may also be equipped with several GIs, prophages, and plasmids. The GIs of *V. cholerae* can excise from the chromosome using tyrosine recombinases and form circular intermediate (20). The circular intermediate of GIs can be lost during cell division, or it may reintegrate in the same locus by site-specific recombination. Although it is well established that the major drivers for genomic plasticity in *V. cholerae* and other enteric pathogens are GIs and other MGEs, their frequency of loss under in vitro and in vivo conditions have not been well studied. In addition, cross-talk among GIs and between GI and prophages are known, but the influence of GI-encoded functions in the essentiality of genes located in the core genome has not been reported. Nevertheless, the molecular basis of host immunity to prevent reinfections of CTXΦ in *V. cholerae* is yet to be understood. In the present study, we have done extensive genome engineering of clinical *V. cholerae* isolates N16961 and 41081 and generated multiple (n = 12) reporter strains to study the in vitro and in vivo stability and integration efficiency of GIs, ICE, and prophages. Each element was tagged with selectable (*cat*/*aadA*), counter-selectable (*sacB*), and chromogenic (*lacZ*) markers to measure the frequency of loss of tagged MGEs under in vitro and in vivo conditions and to isolate different derivatives that lost the GIs and ICE. The frequency of loss of GIs differs widely from 3.26×10^{-4} to 4.21×10^{-5}. GIs that are highly conserved in O1 El Tor, classical, and O139 clinical isolates and essential for virulence are more stable compared to the GIs present in most of the seventh pandemic isolates. Several factors, including mode of integration, expression level of recombinases, and recombination directionality factors, can influence the excision frequencies (40), which is crucial for determining the frequency of loss. It is important to note that, while integration of any GI is reversible, none of these GIs is able to replicate in its excised circular form. After excision, GI may reintegrate into chromosome; otherwise, the excised form cannot be distributed in dividing daughter cells. In sharp contrast, this is not the case for ICE, which has the capacity to excise, replicate, and reintegrate. The frequency of loss of ICE is surprisingly low because of its replication ability in the excised form. Since the integration of CTXΦ is irreversible, we were unable to find any isolate that has lost the prophages.

The simultaneous removal of all of the GIs and prophages from the chromosome of N16961 helped us to understand the cross-talk between core and acquired genomes and importance of horizontally acquired genetic elements in the adaptation of *V. cholerae.* The core genome encodes all of the essential functions of bacteria and is highly conserved within and between bacterial species. The core genome of *V. cholerae* harbors 343 essential genes, while the core genome of *E. coli* houses 287 genes essential for growth in nutrient-rich medium (41, 42). Exclusive essentiality of few genes or domains of genes, including *rctB*, *dam*, *parA2*, and *parB2* in *V. cholerae*, are linked with the replication and segregation of bipartite genome of the cholera pathogen. The SOS master regulator LexA is essential in *V. cholerae* but dispensable in *E. coli* and other bacteria like *Salmonella enterica* and *Shigella dysenteriae*. The reason for the essentiality of LexA for viability of *V. cholerae* cells was not known previously. In this study, we conclusively established that the catalytic function of RstA, a Rep protein essential for the replication of CTXΦ, is the main player to make LexA essential for viability in the toxigenic *V. cholerae* strains. Since CTXΦ and RS1Φ are not present in the genome of *E. coli*, *S. enterica*, *S. dysenteriae*, and other bacterial species, deletion of *lexA* from the genome of these bacteria is feasible. This indicates that, other than encoding CT, tandem copies of CTX prophages also contribute to the fitness of *V. cholerae* by preventing loss or inactivation of such an important molecule in the genome of cholera pathogen. On the other hand, *V. cholerae* cells help CTXΦ for tandem integration, possibly through multiple events of

infections, by providing RecA function. RecA may also help CTXΦ replication by inducing autoproteolysis of LexA in the presence of ssDNA and limiting the availability of RstR, leading to the derepression of the essential phage replication protein RstA (*SI Appendix*, Fig. S3). The cross-biotype CTXΦ infections and their efficient replication in the presence and absence of RecA further support our hypothesis. The major differences in genome sequences between different CTXΦs are mainly located in the *rstR* gene (43). The nucleotide sequences of $rstA^{Cl}$ and $rstB^{Cl}$ are very similar (94% identical) to their El Tor homologs, and the amino acid sequences of these proteins are 99% identical. In contrast, the nucleotide sequences of the *rstR* genes are highly divergent between the two biotypes and other serogroups of *V. cholerae*. The $RstR^{El}$ and $RstR^{Cl}$ share only 24% identical and 32% similar amino acid sequences (43). Such heterogeneity among regulatory proteins and complex interactions between the bacterial core genome and phage-encoded functions possibly facilitates CTXΦ to select compatible host cells for survival and multiplication.

Other than providing toxin and antimicrobial resistance functions, the acquired genome of *V. cholerae*, including GIs, ICE, and prophages, also contributes in modulating host cellular proteome. Core genome-encoded gene functions like LexA can modulate the expression of different genes present in the horizontally acquired genetic elements (44). Chromosomally encoded tyrosine recombinases (XerC and XerD) also contribute substantially in the biology of IMEXs (18). Cross-talk between different acquired genetic elements has also been demonstrated in *V. cholerae* (40). Several VPI-1–encoded proteins modulate the expression of *ctxAB* genes in the genome of CTXΦ under optimal growth conditions and also in the host-associated infection period. Our findings on the differential expression of cellular proteins in the presence and absence of acquired genome provided further insights about their cooperative interactions. In this study, using whole-genome proteome analysis, we have generated further information by identifying an additional pool of proteins ($n = 72$) that are down-regulated in the absence of GIs and prophages. We observed that core genome-encoded proteins like outer membrane proteins ($n = 2$), flagellar biosynthetic proteins ($n = 2$), amino acid modifying enzymes ($n = 2$), and several others are down-regulated in the absence of acquired genetic elements. Since several proteins are involved in various metabolic functions, changes in their expression patterns may also have effects on the cellular metabolome and the fitness of *V. cholerae*.

Conclusion

Our study provides deep insights into the stability of GIs and other MGEs in the genome of *V. cholerae.* These findings will help us to understand how the cholera pathogen has evolved by capturing genetic elements equipped with toxin genes, virulence traits, AMR genes, and other metabolic functions. The findings suggest that the degree of stability of GIs and other MGEs in the genome is directly linked with its importance in pathogenicity and fitness potential of *V. cholerae*. Since CTXΦ encodes the key virulence factor CT, its stability was very high in the genome of *V. cholerae*. The GIs that are important for disease severity and sustainability of *V. cholerae* but not essential for manifesting the disease cholera are less stable compared to CTXΦ. Since gene content variation in the flexible gene pool accumulates more quickly than accumulation of point mutations in the core genome, MGEs and HGT have major contributions in the rapid evolution and emergence of enteric bacterial pathogens. The functions encoded by the core and acquired genome evolved mutually to modulate their essentiality, level of cellular metabolites, fitness of the pathogen, epidemiology, etiology, and sustainability of the *V. cholerae* in host intestine and natural environment. Our findings regarding the essentiality of LexA in toxigenic *V. cholerae* and contribution of RecA in reinfection of CTXΦ established how the core and flexible genome work together to maintain the sustainability of functions that are crucial for virulence, genome repair, and dynamics of GIs and other MGEs. A better understanding of the functions of the genetic contents of MGEs present in the genome of *V. cholerae* and their contribution in the modulation of the proteome would certainly be helpful to forecast evolution and the emergence of enteric pathogens.

Methods

Bacterial Strains, Plasmids, and Culture Conditions. Bacterial strains and plasmids used in this study and their relevant genotypes are listed in Table 1 and *SI Appendix*, Table S1, respectively. *E. coli* and *V. cholerae* cells were grown in Luria broth (LB) with shaking (180 rpm) at 37 °C or Luria agar (LA) unless stated otherwise. Sucrose sensitivity assay was performed by plating the bacteria onto LA supplemented with 15% sucrose and incubating at 24 °C. TCBS (thiosulfate citrate bile salts sucrose) agar was used as a selective medium for *V. cholerae.* Antibiotics were used at the following concentrations: ampicillin (100 μg/mL), streptomycin (100 μg/mL), zeocin (25 μg/mL), kanamycin (40 μg/mL), spectinomycin (50 μg/mL), and chloramphenicol (30 μg/mL for *E. coli* and 2 μg/mL for *V. cholerae*). Blue-white screening was performed by adding 20 μg/mL of X-gal (5-bromo-4-chloro-3-indolyl-β-D-galactopyranoside) on the selection plate. For long-term storage, bacterial strains were maintained at −80 °C in LB containing 20% glycerol (vol/vol).

Transformation and Conjugation. For chemical transformation, rubidium chloride-treated *E. coli* cells were used and subjected to heat shock at 42 °C for 60 s. Conjugation was performed between diaminopimelic acid (DAP) auxotroph *E. coli* β2163 donor cells carrying desired plasmid and *V. cholerae* as recipient cells. Exponentially grown donor and recipient cells were mixed together at a 1:2 ratio, inoculated onto a 0.22-μm sterile filter paper placed on an LA plate supplemented with 0.3 mM DAP, and incubated at 37 °C overnight. Transconjugants were selected based on antibiotic-resistance traits of conjugative plasmids and DAP auxotrophy of donor cells.

Genomic Island Loss Assays. *V. cholerae* strains carrying genomic island tagged with *sacB-cat* or *sacB-cat-lacZ* or *sacB-aadA* were grown to stationary phase in LB at 37 °C with shaking overnight and then plated on LA (15% sucrose) plates supplemented with or without 20 μg/mL of X-gal and incubated for 16 h at 24 °C. About 300 colonies were restreaked on separate LA plates supplemented with chloramphenicol/streptomycin/spectinomycin or sucrose. Subsets of the colonies, which showed sensitivity to chloramphenicol or spectinomycin and resistance to streptomycin and sucrose, were selected for further confirmation by PCR amplification of the genomic island region to be excised.

Genetic Manipulations. The recombinant vectors were developed by using restriction enzymes and T_4 DNA ligase (NEB). Desired DNA fragments were PCR-amplified by using high-fidelity Taq DNA polymerase (NEB). *E. coli* FCV14 was used as host bacterium to select and amplify recombinant vectors. All gene deletions and replacements were performed via allelic exchange method using the derivatives of suicide vector pKAS32 or pDS132 as described previously (26). For delivering vectors to the *V. cholerae*, we adopted a conjugation method. *E. coli* β2163 was used as a donor strain. For all gene deletions, around 500- to 700-base pair (bp) homologous regions upstream and downstream of the respective ORF were PCR-amplified using specific primer combinations (*SI Appendix*, Table S3). The amplified products were purified before restriction enzyme digestion and introduced into similarly digested suicide vector by DNA ligation.

For removing the entire genomic islands (GIs) from *V. cholerae* genome, first, the GIs were tagged with *sacB-cat*, *sacB-aadA*, or *sacB-cat-lacZ* alleles. Briefly, a portion of the GI was PCR-amplified and then cloned into pSW23T (*SI Appendix*, Table S1). The recombinant vectors were used as a template for inverse PCR by which a portion GI was replaced with *sacB-cat*, *sacB-aadA*, or *sacB-cat-lacZ* was cloned. Finally, the recombinant vectors carrying *sacB-cat*, *sacB-aadA*, or *sacB-cat-lacZ* alleles flanked by the small segments of GIs were used to tag the GIs by the sequential allelic exchange method as described previously (45).

The GI flagged strains became resistant to chloramphenicol/spectinomycin but sensitive to sucrose. For the removal of the entire GI, the strains carrying flagged GI were grown in LB without any antibiotics. The *V. cholerae* strain that lost its GI was isolated simply by plating the *sacB-cat*, *sacB-aadA*, or *sacB-cat-lacZ* flagged strain on to LA plates supplemented with 15% sucrose with or without X-gal. The loss of GIs from the bacterium was confirmed by monitoring sensitivity to chloramphenicol/spectinomycin and also by PCR assay.

Rabbit Ileal Loop Assay. Rabbit ileal loop assay was performed in young New Zealand White rabbits (1.5 to 2.0 kg) as described earlier (46). Prior to surgery, rabbits were fasted for 24 h, but water was given ad libitum. *V. cholerae* cells were grown in LB up to mid-log phase, followed by washing with sterile phosphate-buffered saline (PBS). Rabbits were anesthetized, laparotomy was performed, and the ileum was ligated into discrete loops, each measuring ~10 cm. Each loop was inoculated with 10^8 CFU of *V. cholerae* strain (N16961, BS1, SB31, or SB55) in PBS. The loop containing only PBS was used as a negative control. The intestine was returned to the peritoneum, and the abdomen was sutured. Rabbits were euthanized after 18 h, and the abdomen was opened, followed by measurement of loop length and accumulated fluid volume in each loop. Fluid accumulation (FA) ratio was calculated by the following formula: volume of fluid accumulated (milliliters)/length of intestine (centimeters). FA ratio of greater than 1.0 was considered as a strong positive response, while FA ratio of less than 0.2 indicated negative response. All the animal experiments were conducted following the standard operating procedure as outlined by the Committee for the Purpose of Control and Supervision of Experiments on Animals (CPCSEA), Ministry of Environment and Forest, Government of India. The animal experimental protocol was approved by the Institutional Animal Ethical Committee (IAEC) of National Institute of Cholera and Enteric Diseases (NICED) with Registration No. 68//Rebi/S/1999/CPCSEA valid 17/7/2024, Approval No.: NICED/CPCSEA/68/GO/(25/294)/2019-IAEC/SM/1.

To measure the in vivo stability of the pathogenicity islands of *V. cholerae*, the *sacB-cat*–tagged strains (SB23, SB26, SB30, and DD1) were inoculated into rabbit ileum, and then the animal passaged strains were plated onto both TCBS agar and LA supplemented with 15% sucrose and streptomycin (100 μg/mL).

Whole-Cell Proteome Analysis. Overnight-grown *V. cholerae* strains N16961 and BB1 were used for total protein extraction and trypsin digestion. We used an LC system (Eksigent 2D) coupled with a Triple TOF 5600^+ mass spectrometer (AB Sciex) for proteome analysis. The *V. cholerae*-specific library was created by bRP-C_{18}-HPLC followed by LC-MS/MS for label-free quantitation in SWATH-MS workflow. In brief, the peptides from the cell lysate of both the *V. cholerae* strains were fractionated using an Agilent 1200 HPLC system on a high-pH reverse-phase Zorbax 300 Extend-C18 column (Agilent Technology). The eluted 48 fractions were pooled to the final 12 fractions. Each fraction was desalted and subjected to 90-min gradient LC-MS/MS run in data-dependent acquisition (DDA) mode to generate a *V. cholerae* spectral library. For quantitative analysis, each sample was acquired in triplicate in data-independent acquisition (DIA) mode and the instrument was configured as described by Gillet et al. (47). MaxQuant version 1.6 was used for peptide and protein identification, and a Spectronaut pulsar (Biognosys) with default settings was used in SWATH processing and quantification. Proteins with a ratio of >1.5 were considered differentially expressed. In all cases, $P < 0.05$ (*t* test) was considered significant in protein quantification.

Data Availability. All study data are included in the article and *SI Appendix*.

ACKNOWLEDGMENTS. We thank Dr. Raghunath Chatterjee, Human Genetics Unit, Indian Statistical Institute, Kolkata, India, for critical reading of the manuscript. The authors acknowledge Mr. Naveen Kumar and other lab members for technical support. This study received financial support from the Department of Biotechnology (DBT), Government of India (Grant No. BT/MB/THSTI/HMC-SFC/2011). A.P. and J.V. received support from the DBT, Government of India. P.K. received research fellowship from the Council of Scientific and Industrial Research (CSIR), Government of India.

1. S. Casjens, The diverse and dynamic structure of bacterial genomes. *Annu. Rev. Genet.* **32**, 339–377 (1998).
2. M. K. Waldor, J. J. Mekalanos, Lysogenic conversion by a filamentous phage encoding cholera toxin. *Science* **272**, 1910–1914 (1996).
3. D. K. Karaolis, S. Somara, D. R. Maneval, Jr., J. A. Johnson, J. B. Kaper, A bacteriophage encoding a pathogenicity island, a type-IV pilus and a phage receptor in cholera bacteria. *Nature* **399**, 375–379 (1999).
4. L. S. Frost, R. Leplae, A. O. Summers, A. Toussaint, Mobile genetic elements: The agents of open source evolution. *Nat. Rev. Microbiol.* **3**, 722–732 (2005).
5. J. Verma *et al*., Genomic plasticity associated with antimicrobial resistance in *Vibrio cholerae*. *Proc. Natl. Acad. Sci. U.S.A.* **116**, 6226–6231 (2019).
6. P. Spanogiannopoulos, E. N. Bess, R. N. Carmody, P. J. Turnbaugh, The microbial pharmacists within us: A metagenomic view of xenobiotic metabolism. *Nat. Rev. Microbiol.* **14**, 273–287 (2016).
7. B. S. Samuel *et al*., Genomic and metabolic adaptations of Methanobrevibacter smithii to the human gut. *Proc. Natl. Acad. Sci. U.S.A.* **104**, 10643–10648 (2007).
8. U. Dobrindt, B. Hochhut, U. Hentschel, J. Hacker, Genomic islands in pathogenic and environmental microorganisms. *Nat. Rev. Microbiol.* **2**, 414–424 (2004).
9. M. G. Langille, W. W. Hsiao, F. S. Brinkman, Detecting genomic islands using bioinformatics approaches. *Nat. Rev. Microbiol.* **8**, 373–382 (2010).
10. E. F. Boyd, S. Almagro-Moreno, M. A. Parent, Genomic islands are dynamic, ancient integrative elements in bacterial evolution. *Trends Microbiol.* **17**, 47–53 (2009).
11. B. Das, Insights into TLCΦ lysogeny: A twist in the mechanism of IMEX integration. *Proc. Natl. Acad. Sci. U.S.A.* **116**, 18159–18161 (2019).
12. J. Hacker, U. Hentschel, U. Dobrindt, Prokaryotic chromosomes and disease. *Science* **301**, 790–793 (2003).
13. J. F. Heidelberg *et al*., DNA sequence of both chromosomes of the cholera pathogen *Vibrio cholerae*. *Nature* **406**, 477–483 (2000).
14. F. X. Weill *et al*., Genomic history of the seventh pandemic of cholera in Africa. *Science* **358**, 785–789 (2017).
15. D. Domman *et al*., Integrated view of *Vibrio cholerae* in the Americas. *Science* **358**, 789–793 (2017).
16. A. Mutreja *et al*., Evidence for several waves of global transmission in the seventh cholera pandemic. *Nature* **477**, 462–465 (2011).
17. N. D. Grindley, K. L. Whiteson, P. A. Rice, Mechanisms of site-specific recombination. *Annu. Rev. Biochem.* **75**, 567–605 (2006).
18. B. Das, E. Martínez, C. Midonet, F. X. Barre, Integrative mobile elements exploiting Xer recombination. *Trends Microbiol.* **21**, 23–30 (2013).
19. C. Rajanna *et al*., The vibrio pathogenicity island of epidemic Vibrio cholerae forms precise extrachromosomal circular excision products. *J. Bacteriol.* **185**, 6893–6901 (2003).
20. R. A. Murphy, E. F. Boyd, Three pathogenicity islands of Vibrio cholerae can excise from the chromosome and form circular intermediates. *J. Bacteriol.* **190**, 636–647 (2008).
21. A. M. Devault *et al*., Second-pandemic strain of Vibrio cholerae from the Philadelphia cholera outbreak of 1849. *N. Engl. J. Med.* **370**, 334–340 (2014).
22. M. K. Waldor, H. Tschäpe, J. J. Mekalanos, A new type of conjugative transposon encodes resistance to sulfamethoxazole, trimethoprim, and streptomycin in Vibrio cholerae O139. *J. Bacteriol.* **178**, 4157–4165 (1996).
23. C. J. Grim *et al*., Genome sequence of hybrid Vibrio cholerae O1 MJ-1236, B-33, and CIRS101 and comparative genomics with V. cholerae. *J. Bacteriol.* **192**, 3524–3533 (2010).
24. M. E. Val *et al*., The single-stranded genome of phage CTX is the form used for integration into the genome of Vibrio cholerae. *Mol. Cell* **19**, 559–566 (2005).
25. B. Das, J. Bischerour, M. E. Val, F. X. Barre, Molecular keys of the tropism of integration of the cholera toxin phage. *Proc. Natl. Acad. Sci. U.S.A.* **107**, 4377–4382 (2010).
26. A. Pant *et al*., Effect of LexA on chromosomal integration of CTXϕ in Vibrio cholerae. *J. Bacteriol.* **198**, 268–275 (2015).
27. S. Pradhan, A. K. Baidya, A. Ghosh, K. Paul, R. Chowdhury, The El Tor biotype of Vibrio cholerae exhibits a growth advantage in the stationary phase in mixed cultures with the classical biotype. *J. Bacteriol.* **192**, 955–963 (2010).
28. A. Saha, S. Haralalka, R. K. Bhadra, A naturally occurring point mutation in the 13-mer R repeat affects the oriC function of the large chromosome of *Vibrio cholerae* O1 classical biotype. *Arch. Microbiol.* **182**, 421–427 (2004).
29. S. M. Faruque, J. J. Mekalanos, Pathogenicity islands and phages in Vibrio cholerae evolution. *Trends Microbiol.* **11**, 505–510 (2003).
30. E. Martínez, J. Campos-Gómez, F. X. Barre, CTXϕ: Exploring new alternatives in host factor-mediated filamentous phage replications. *Bacteriophage* **6**, e1128512 (2016).
31. H. H. Kimsey, M. K. Waldor, Vibrio cholerae LexA coordinates CTX prophage gene expression. *J. Bacteriol.* **191**, 6788–6795 (2009).
32. H. Ochman, L. M. Davalos, The nature and dynamics of bacterial genomes. *Science* **311**, 1730–1733 (2006).
33. F. X. Weill *et al*., Genomic insights into the 2016-2017 cholera epidemic in Yemen. *Nature* **565**, 230–233 (2019).
34. K. Okada, T. Iida, K. Kita-Tsukamoto, T. Honda, Vibrios commonly possess two chromosomes. *J. Bacteriol.* **187**, 752–757 (2005).
35. M. Bruhn *et al*., Functionality of two origins of replication in *Vibrio cholerae* strains with a single chromosome. *Front. Microbiol.* **9**, 2932 (2018).

36. J. W. Schroeder, P. Yeesin, L. A. Simmons, J. D. Wang, Sources of spontaneous mutagenesis in bacteria. *Crit. Rev. Biochem. Mol. Biol.* **53**, 29–48 (2018).
37. I. Levade *et al*., Vibrio cholerae genomic diversity within and between patients. *Microb. Genom.* **3**, e000142 (2017).
38. H. Ochma, I. B. Jones, Evolutionary dynamics of full genome content in Escherichia coli. *EMBO J.* **19**, 6637–6643 (2000).
39. S. M. Faruque *et al*., Genetic diversity and virulence potential of environmental Vibrio cholerae population in a cholera-endemic area. *Proc. Natl. Acad. Sci. U.S.A.* **101**, 2123–2128 (2004).
40. M. R. Carpenter, S. Rozovsky, E. F. Boyd, Pathogenicity island cross talk mediated by recombination directionality factors facilitates excision from the chromosome. *J. Bacteriol.* **198**, 766–776 (2015).
41. M. C. Chao *et al*., High-resolution definition of the Vibrio cholerae essential gene set with hidden Markov model-based analyses of transposon-insertion sequencing data. *Nucleic Acids Res.* **41**, 9033–9048 (2013).
42. J. Kato, M. Hashimoto, Construction of consecutive deletions of the Escherichia coli chromosome. *Mol. Syst. Biol.* **3**, 132 (2007).
43. H. H. Kimsey, M. K. Waldor, CTXphi immunity: Application in the development of cholera vaccines. *Proc. Natl. Acad. Sci. U.S.A.* **95**, 7035–7039 (1998).
44. E. Krin *et al*., Expansion of the SOS regulon of Vibrio cholerae through extensive transcriptome analysis and experimental validation. *BMC Genomics* **19**, 373 (2018).
45. A. Kumar, S. Bag, B. Das, Novel genetic tool to study the stability of genomic islands. *Recent Pat. Biotechnol.* **12**, 200–207 (2018).
46. S. N. De, D. N. Chatterje, An experimental study of the mechanism of action of Vibriod cholerae on the intestinal mucous membrane. *J. Pathol. Bacteriol.* **66**, 559–562 (1953).
47. L. C. Gillet *et al*., Targeted data extraction of the MS/MS spectra generated by data-independent acquisition: A new concept for consistent and accurate proteome analysis. *Mol. Cell. Proteomics* **11**, O111.016717 (2012).