RESEARCH ARTICLE

# Acquisition and Evolution of SXT-R391 Integrative Conjugative Elements in the Seventh-Pandemic *Vibrio cholerae* Lineage

**Matteo Spagnoletti,[a] Daniela Ceccarelli,[b] Adrien Rieux,[a] Marco Fondi,[c] Elisa Taviani,[d,e] Renato Fani,[c] Mauro M. Colombo,[d,e] Rita R. Colwell,[b,f] François Balloux[a]**

University College London Genetics Institute, University College London, London, United Kingdom[a]; Maryland Pathogen Research Institute, University of Maryland, College Park, Maryland, USA[b]; Department of Biology, University of Florence, Florence, Italy[c]; Department of Biology and Biotechnology Charles Darwin, Sapienza University of Rome, Rome, Italy[d]; Centro de Biotecnologia, Universitade Eduardo Mondlane, Maputo, Mozambique[e]; Johns Hopkins Bloomberg School of Public Health, Johns Hopkins University, Baltimore, Maryland, USA[f]

**ABSTRACT** **SXT-R391 Integrative conjugative elements (ICEs) are self-transmissible mobile genetic elements able to confer multidrug resistance and other adaptive features to bacterial hosts, including *Vibrio cholerae*, the causative agent of cholera. ICEs are arranged in a mosaic genetic structure composed of a conserved backbone interspersed with variable DNA clusters located in conserved hot spots. In this study, we investigated ICE acquisition and subsequent microevolution in pandemic *V. cholerae*. Ninety-six ICEs were retrieved from publicly available sequence databases from *V. cholerae* clinical strains and were compared to a set of reference ICEs. Comparative genomics highlighted the existence of five main ICE groups with a distinct genetic makeup, exemplified by ICE*Vch*Ind5, ICE*Vch*Moz10, SXT, ICE*Vch*Ind6, and ICE*Vch*Ban11. ICE*Vch*Ind5 (the most frequent element, represented by 70 of 96 elements analyzed) displayed no sequence rearrangements and was characterized by 46 single nucleotide polymorphisms (SNPs). SNP analysis revealed that recent inter-ICE homologous recombination between ICE*Vch*Ind5 and other ICEs circulating in gammaproteobacteria generated ICE*Vch*Moz10, ICE*Vch*Ind6, and ICE*Vch*Ban11. Bayesian phylogenetic analyses indicated that ICE*Vch*Ind5 and SXT were independently acquired by the current pandemic *V. cholerae* O1 and O139 lineages, respectively, within a period of only a few years.**

**IMPORTANCE** **SXT-R391 ICEs have been recognized as key vectors of antibiotic resistance in the seventh-pandemic lineage of *V. cholerae*, which remains a major cause of mortality and morbidity on a global scale. ICEs were acquired only recently in this clade and are acknowledged to be major contributors to horizontal gene transfer and the acquisition of new traits in bacterial species. We have reconstructed the temporal dynamics of SXT-R391 ICE acquisition and spread and have identified subsequent recombination events generating significant diversity in ICEs currently circulating among *V. cholerae* clinical strains. Our results showed that acquisition of SXT-R391 ICEs provided the *V. cholerae* seventh-pandemic lineage not only with a multidrug resistance phenotype but also with a powerful molecular tool for rapidly accessing the pan-genome of a large number of gammaproteobacteria.**

**Received** 21 May 2014 **Accepted** 11 July 2014 **Published** 19 August 2014

**Citation** Spagnoletti M, Ceccarelli D, Rieux A, Fondi M, Taviani E, Fani R, Colombo MM, Colwell RR, Balloux F. 2014. Acquisition and evolution of SXT-R391 integrative conjugative elements in the seventh-pandemic *Vibrio cholerae* lineage. mBio 5(4):e01356-14. doi:10.1128/mBio.01356-14.

**Editor** Julian Davies, University of British Columbia



Address correspondence to Matteo Spagnoletti, m.spagnoletti@ucl.ac.uk.

*Vibrio cholerae*, the causative agent of cholera, is a native inhabitant of the aquatic ecosystem and represents a serious issue for global public health in developing countries. More than 200 serogroups characterized by different somatic O antigens have been described to date, but only serogroups O1 and O139 have been linked to epidemics (1). Seven cholera pandemics have been recorded since 1817 (2). The current seventh pandemic started in 1961 and is caused by *V. cholerae* O1 biotype El Tor, which replaced the Classical biotype responsible for previous pandemics (1). Several sublineages of *V. cholerae* O1 El Tor have emerged over the past 50 years, the most noteworthy being *V. cholerae* O139 Bengal (3, 4). This epidemic serogroup emerged in late 1992 in the Indian subcontinent, and its distribution is currently limited to Asia (3, 5).

Variants of *V. cholerae* O1 showing features of both Classical and El Tor biotypes have been repeatedly isolated in Asia and Africa and are collectively referred to as atypical El Tor (6). In recent years, several studies using comparative genomics have elucidated the evolutionary relationships of these genetically distinct *V. cholerae* variants (5, 7–9). It was demonstrated that they all belong to a single lineage known as the current seventh-pandemic clade (5). Mutreja et al. produced a robust phylogenetic reconstruction for the lineage responsible for the seventh pandemic, suggesting that it spread from the Bay of Bengal in at least three overlapping waves of strains that shared a common ancestor in the 1950s (8). An event at the point of transition between the wave 1 and wave 2 pandemic clones was the acquisition of an integrative conjugative element (ICE) of the SXT-R391 family (8, 10). This transition was dated to approximately 1978 to 1984 and coincides with the dating of the most recent common ancestor (MRCA) of

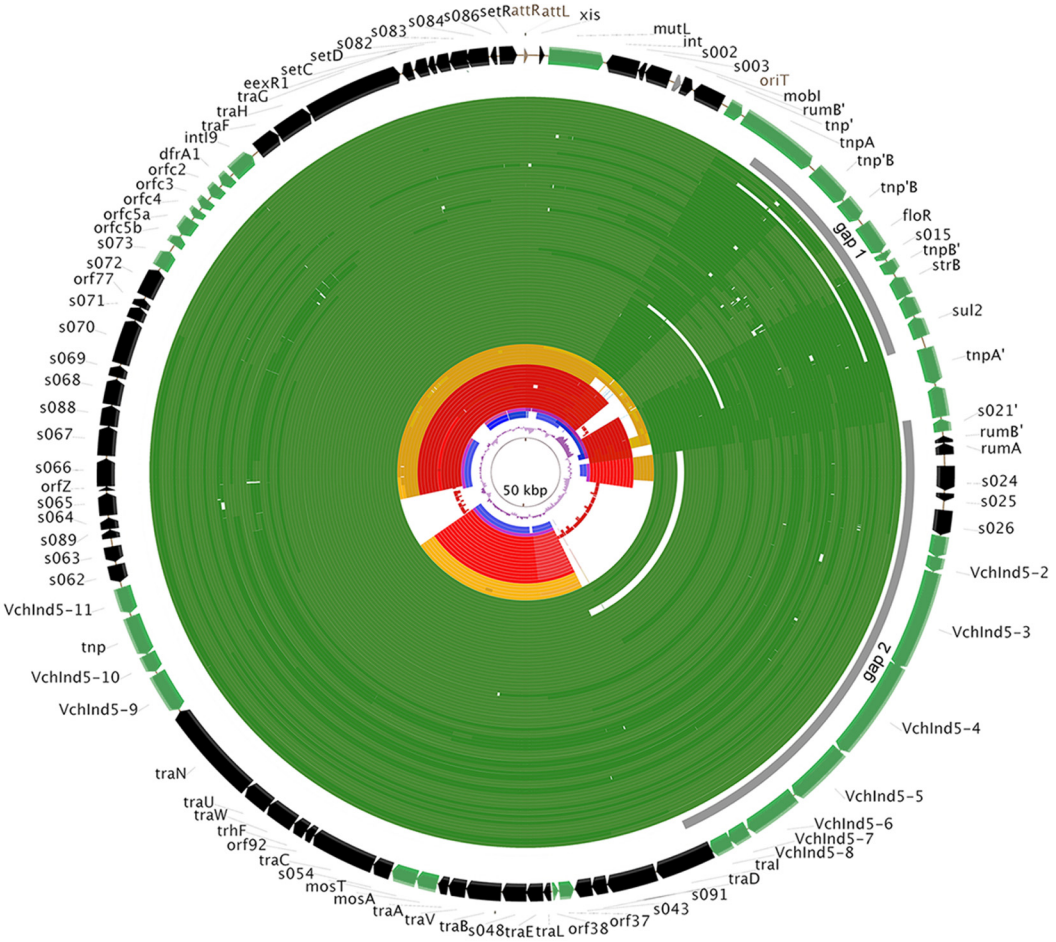


**FIG 1** BLAST atlas using ICE*Vch*Ind5 showing the genetic rearrangements in the ICE data set. Contigs from each *de novo* assembly of 96 *V. cholerae* genomes were mapped against the sequence of ICE*Vch*Ind5 (outer circle, with annotations). ICE profiles are shown with different colors according to their variable levels of genetic content within the hot spots. Green, ICE*Vch*Ind5 (74 ICEs); orange, ICE*Vch*Moz10 (6 ICEs); red, ICE*Vch*Ind6 (13 ICEs); purple, ICE*Vch*Ban11 (1 ICE); blue, SXT (2 ICEs). Absence of color represents missing genetic material compared to ICE*Vch*Ind5. Gray bars represent deletion regions gap 1 and gap 2 in the antibiotic resistance gene cluster and hotspot 5, respectively. Color gradients are proportional to the BLAST percent identity (90% to 100%).

the O1 and O139 serogroups. None of the *V. cholerae* O1 strains belonging to the first wave of pandemic transmission contain an ICE (8).

SXT-R391 ICEs are a well-studied family of self-transmissible mobile genetic elements, and recent findings suggest that these elements are present in a wide range of environmental *Gammaproteobacteria*, occupying different ecological niches in the aquatic environment (11, 12). The prototypical elements, SXT and R391, were isolated from *V. cholerae* O139 in India and *Providencia rettgeri* in South Africa, respectively (13, 14). ICEs are able to transfer by conjugation, using a circular intermediate, and integrate into and replicate as part of the host chromosome (15). The backbone of ICEs is composed of 52 core genes, of which 25 are required for key functions of integration/excision, conjugative transfer, and regulation (10). This backbone serves as a scaffold for the integration of variable DNA located within conserved sites named variable regions (VR1, VR2, and VR3) and "hotspots" (HS1, HS2, HS3, HS4, and HS5) (10). Core genes are ~50 kb in overall size, and an additional variable genetic cargo of ~30 to ~60 kb can be inserted into the hot spots (10). Variable DNA content can confer ICE-specific features, such as multidrug and heavy-metal resistance, restriction modification systems, and alternative metabolic pathways, to the bacterial host (12, 16–18). After the identification of the SXT element in *V. cholerae* O139 (13), comparative analysis revealed that at least two ICEs of the same family, ICE*Vch*Ind5 and ICE*Vch*Moz10, were circulating in *V. cholerae* O1 El Tor strains and that both showed rearrangements in their variable regions compared with SXT (10, 15). ICE*Vch*Ind5 was recognized as the prevalent ICE associated with *V. cholerae* El Tor strains circulating worldwide (19–21), but a wide-range study using genomic data is still lacking.

Recombination appears to play a crucial role in shaping ICE structure and driving its evolution (22, 23). Inter-ICE homologous recombination can take place through the exchange of large DNA fragments between two different ICEs redundantly transferred in the same cell (22). Such recombination happens frequently in *in vitro* experiments and depends on host (*recA*) and ICE (*bet* and *exo*) genes (22, 24). Subsequent conjugative transfer of two ICEs can occur, as only two exclusion groups (S and R) have been detected within the family (25). ICEs of different exclusion groups can temporarily coexist in the same host chromosome with formation of tandem arrays (23). Arrays are excellent sub-

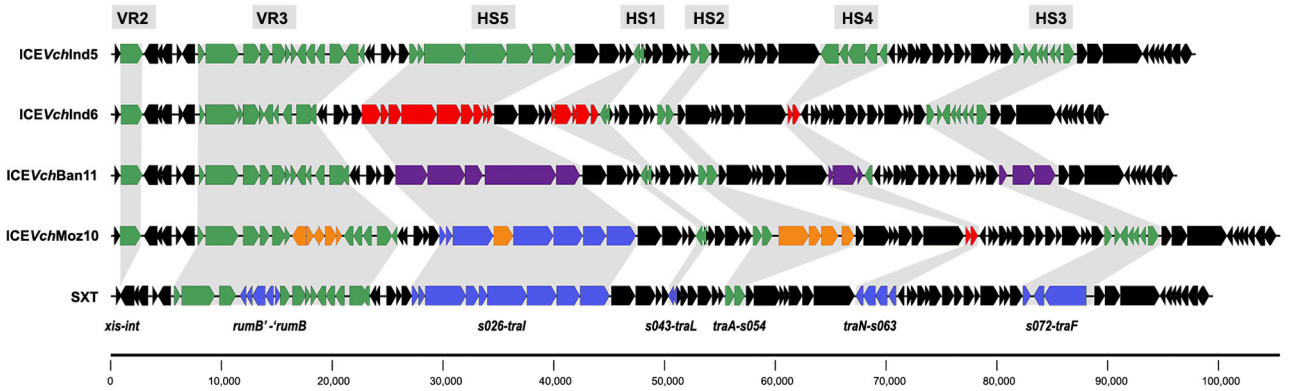


**FIG 2** Genetic organization of the five ICE profiles detected in *V. cholerae* clinical strains. Backbone genes are depicted in black. Different hot spots and variable regions are highlighted in gray. Variable genes shared between ICE*Vch*Ind5 and other ICEs are represented in green. Genes present in both SXT and ICE*Vch*Moz10 are depicted in blue. Genes present in both ICE*Vch*Ind6 and ICE*Vch*Moz10 are depicted in red. Genes unique to ICE*Vch*Ban11 and ICE*Vch*Moz10 are represented in purple and orange, respectively.

strates for recombination and promote formation of hybrid ICEs, while backbone synteny is ensured by selection of functional ICEs able to be transferred and maintained in the bacterial population (22). Transposases, ISCR (insertion sequence common region) elements, and other insertion sequences are abundant among variable regions of SXT-R391 ICEs (10, 26); they can promote genetic rearrangements and/or acquisition of new cargo genes in the ICE genetic scaffold (27). A remarkable example of this phenomenon is provided by the organization of the SXT resistance cluster, a ~14-kb region composed of more than seven transposases and ISCR*2* elements, which embeds multiple resistance determinants between *rumB* and *rumA* (26, 28).

ICE genetic organization, transfer, and regulation have been extensively studied (15, 29), but little is known about the ICE evolutionary dynamics of a single host species. Here we report the prevalence, diversity, and dynamics of recombination within the SXT-R391 family, as well as their appearance and spread in the seventh-pandemic *V. cholerae* lineage. Comparative and recombination analyses were complemented with an estimate of ICE acquisition in order to reconstruct a comprehensive picture of processes underlying microevolution of this family of mobile genetic elements.

## RESULTS

**SXT-R391 ICEs in *Vibrio cholerae* O1 and O139 clinical strains.** Selection of *V. cholerae* genomes was carried out based on the year of strain isolation, geographic location, nucleotide sequence quality, and presence of an ICE (see Table S1 in the supplemental material). The overall ICE genetic organization in the data set was investigated to detect structural rearrangements within core and variable regions. All *V. cholerae* genomes analyzed (see Table S1) were *de novo* assembled and the assemblies screened against reference elements ICE*Vch*Ind5, ICE*Vch*Moz10, and SXT, prevalent ICEs associated with *V. cholerae* O1 and O139 worldwide (8, 10, 19). The presence or absence of variable genes with respect to reference ICEs was determined using BLAST (30), and BLAST atlases were constructed for each of the three ICEs (Fig. 1; see also Fig. S1A and S1B in the supplemental material). The BLAST atlas showing structural rearrangements of all 96 elements against the most prevalent ICE in the data set (ICE*Vch*Ind5) is presented in Fig. 1.

Eighty-two of 96 ICEs had >98% sequence similarity with elements previously detected in clinical *V. cholerae* isolates as follows: for ICE*Vch*Ind5, profile 1 (see Table S1 in the supplemental material; shown in green in Fig. 1); for ICE*Vch*Moz10, profile 2 (see Table S1; shown in orange in Fig. 1); and for SXT, profile 5 (see Table S1; shown in blue in Fig. 1). The remaining 14 ICEs carried major rearrangements in the variable regions compared to ICE*Vch*Ind5 (innermost circles in red and purple) (Fig. 1). Thirteen of the 14 elements revealed identical genetic contents and were retrieved from *V. cholerae* O1 strains isolated in India, Nepal, and Thailand (8, 31) (profile 3 [see Table S1; shown in red in Fig. 1]). An additional ICE, isolated in Bangladesh in 2000, was unique in its variable DNA content (profile 4 [see Table S1; shown in purple in Fig. 1]). These novel ICEs were named ICE*Vch*Ind6 ($n = 13$) and ICE*Vch*Ban11 ($n = 1$), respectively.

We detected structural variations within ICE groups due to major deletions. For example, six ICEs belonging to ICE*Vch*Ind5 (*V. cholerae* O1, Bangladesh 1991, and Tanzania 2009), ICE*Vch*Moz10 (*V. cholerae* O1, Mozambique 2005), and SXT (*V. cholerae* O139, Bangladesh 2002) (see Table S1 in the supplemental material) shared a large deletion in the antibiotic resistance cluster comprising genes *floR*, *strB*, and *sul2* (gap 1; Fig. 1). A second major rearrangement is represented by gap 2 (Fig. 1). Two ICEs belonging to the ICE*Vch*Ind5 group and isolated in Nepal in 2010 (see Table S1) exhibited a deletion from *rumA* to *traI*. The deletion comprised backbone genes *s024*, *s025*, and *s026*, as well as the entire genetic cluster located in hotspot 5 and part of *traI* encoding a putative relaxase. The two ICEs are likely defective elements unable to transfer by conjugation, as a functional *traI* is required for successful ICE transfer (32).

The major gaps for hotspot 4, hotspot 3, variable region 2, and hotspot 5 observed in the BLAST atlas of ICE*Vch*Ind5 (Fig. 1) are due to the different variable-region content of each ICE. The genetic organization of the five ICE profiles is shown in Fig. 2. Since variable genes encode a large array of functions, only hotspots 3 and 5 are discussed here. Hotspot 3 contains genes that encode diguanylate cyclases in SXT, a class 4 integron carrying the trimethoprim resistance *dfrA1* gene in ICE*Vch*Ind5, ICE*Vch*Ind6, and ICE*Vch*Moz10 and two putative genes annotated as exonucleases and helicases in ICE*Vch*Ban11. Higher variability is

observed in hotspot 5, which carries a cluster of nine genes of unknown function in ICE*Vch*Ind5, a possible restriction modification system in ICE*Vch*Ind6, five open reading frames (ORFs) annotated as helicases in ICE*Vch*Ban11, and a cluster of eight ORFs, including those encoding a putative ATPase, type II restriction enzymes, and a phage growth limitation factor in ICE*Vch*Moz10 and SXT (Fig. 2) (10).

**Genetic variability and signatures of recombination among ICE groups.** To study genetic diversity within and between the five ICE profiles, single nucleotide polymorphisms (SNPs) were called by comparing reads of all isolates to ICE*Vch*Ind5. We discovered a total of 2,967 SNPs among the 96 ICEs. The reads of each ICE belonging to a specific profile were mapped against the reference ICE selected for each group: ICE*Vch*Ind5 ($n = 70$, 46 SNPs), ICE*Vch*Moz10 ($n = 6$, 12 SNPs), SXT ($n = 2$, 8 SNPs), and ICE*Vch*Ind6 ($n = 13$, 4 SNPs). Analysis was not done on ICE*Vch*Ban11 as it was represented by a single isolate. The vast majority of SNPs occurred between the five ICE groups, and only a limited number of SNPs was detected within each ICE group.

To assess and compare the levels of recombination in core versus noncore ICE regions, we built neighbor-net networks, as implemented by SplitsTree 4 (33). Two separate datasets were considered, one comprising the entire ICE sequence and one with only the core backbone. The five *V. cholerae* ICE profiles were used, together with those of all SXT-R391 ICEs available to date that have been isolated from environmental non-O1/non-O139 *V. cholerae* strains and other *Gammaproteobacteria* species (i.e., *Providencia rettgeri*, *Providencia alcalifaciens*, *Shewanella putrefaciens*, *Proteus mirabilis*, *Photobacterium damselae*, and *Vibrio fluvialis*) (see Table S1 in the supplemental material). The resultant neighbor-net networks were characterized by extensive reticulation, as shown by analysis of both whole ICE sequences and core regions only (see Fig. S2), and the results of the Phi test for recombination were highly statistically significant ($P < 0.0001$) in both analyses. The presence of extensive reticulation in the core gene set (see Fig. S2B) points to inter-ICE homologous recombination as an important force shaping ICE structures.

**Evidence of recent inter-ICE homologous recombination in seventh-pandemic *V. cholerae*.** The distribution of SNPs along the 52 conserved core genes of five representative ICEs was further investigated by directly mapping reads from ICE*Vch*Ban11, ICE*Vch*Ind6, ICE*Vch*Moz10, and SXT isolates against the conserved backbone of ICE*Vch*Ind5. To identify possible donors to hybrid ICE formation, reads were also mapped against core genes of the same additional eight ICEs used for network analysis. Potential recent inter-ICE recombination events were identified by detecting dense clusters of SNPs separated by clear-cut boundaries (34).

This analysis indicated that ICE*Vch*Ban11 was generated by recombination between ICE*Vch*Ind5 and ICE*Pmi*USA1, which was isolated from a uropathogenic *Proteus mirabilis* strain (Fig. 3A). Similarly, ICE*Vch*Ind6 underwent hybrid rearrangement between ICE*Vch*Ind5 and ICE*Vfl*Ind1 isolated from a clinical *V. fluvialis* strain. Dense clusters of SNPs between *traD* and *s068* were detected, while the rest of the sequence showed no SNPs compared to ICE*Vch*Ind5 (Fig. 3B). ICE*Vch*Moz10 showed recombination similar to that seen with ICE*Vch*Ind6, sharing large regions of the backbone with ICE*Vch*Ind5 and ICE*Vfl*Ind1 (Fig. 3C). Analysis revealed no SNPs in the first ~7 kb (*xis*-to-*rumB* region) of ICE*Vch*Ind5, ICE*Vch*Ban11, ICE*Vch*Ind6, and ICE*Vch*Moz10 (Fig. 3A to C). Conversely, the distributions of SNPs in SXT and ICE*Vch*Ind5 did not point to evidence of recent homologous recombination affecting backbone genes of the two elements (Fig. 3D).

**Acquisition date of ICE*Vch*Ind5 and SXT.** Acquisition of SXT-R391 ICEs was dated by Mutreja and colleagues as occurring between 1978 and 1984 by the use of genome-wide SNP data within a Bayesian framework (8). Those authors estimated the lower boundary for acquisition of ICEs as the time to the MRCA (TMRCA) of all *V. cholerae* strains carrying ICEs. In this study, we refined the dating analysis by estimating directly the age of ICE*Vch*Ind5 and that of SXT rather than the TMRCA of the strains carrying them. We included ICE sequences from the ICE*Vch*Ind5 and SXT groups, but ICEs shown to be recombinants (ICE*Vch*Moz10, ICE*Vch*Ind6, and ICE*Vch*Ban11) were not considered. Nucleotide alignments were built on ICE*Vch*Ind5-like and SXT-like ICEs separately, in order to reconstruct two phylogenies and date the acquisition of each of the two groups of ICEs independently.

Previous analysis supported identification of ICE*Vch*Ind5 as the most prevalent ICE, i.e., present in 74 of 96 samples isolated between 1989 and 2010, and likely the oldest ICE in our data set. Inferences were realized using BEAST (35) on an alignment of 71 ICEs (70 ICEs plus ICE*Vch*Ind5, excluding four elements with deletions) (see Fig. S3A in the supplemental material). The time of acquisition of ICE*Vch*Ind5 (i.e., the TMRCA) was dated to 1985 (95% highest posterior density [HPD] interval, 1980 to 1989) (Fig. 4A). Posterior distributions of substitution rates are given with confidence intervals in Fig. S4. To ensure that our estimate for the time of acquisition of ICE*Vch*Ind5 elements was not affected by cryptic recombination, we tested for recombination between ICE*Vch*Ind5 elements using five different algorithms implemented in Recombination Detection Program 4 (RDP4) (36) and the Phi test from the PhiPack software (33). While the results of all tests obtained with RDP4 turned out to be statistically nonsignificant, those of the Phi test were borderline significant ($P = 0.039$). Using SplitsTree4 software and manually inspecting the alignments, we created a reduced alignment by removing nine ICEs that were showing potential recombination (see Table S1). This new reduced alignment of 62 ICEs led to nonsignificant Phi test results ($P = 0.302$) and a median TMRCA qualitatively similar to that of the whole alignment (results not shown), indicating that the dating analysis was not affected by cryptic recombination between ICE*Vch*Ind5 elements.

We repeated similar analyses of four sequences of SXT-like ICEs extracted from *V. cholerae* O139 isolates between 1992, the year of its first appearance, and 2002. For this set of ICEs, acquisition was dated to 1992 (95% HPD, 1990 to 1992) (Fig. 4B).

## DISCUSSION

The availability of whole-genome sequence data from *V. cholerae* clinical isolates offers a remarkable opportunity to study the microevolution of SXT-R391 ICEs in a single host species. To date, most of the studies have focused on the evolution of the current pandemic lineage of *V. cholerae*, where SXT-R391 ICEs are ubiquitous (5, 8, 9). Both Chun et al. and Mutreja et al. have hinted at the acquisition of SXT-related ICEs as an important milestone in *V. cholerae* genome dynamics and evolution (5, 8). Our report refines this scenario and offers new insights into the dynamics of acquisition and recombination of the SXT-R391 family of ICEs.

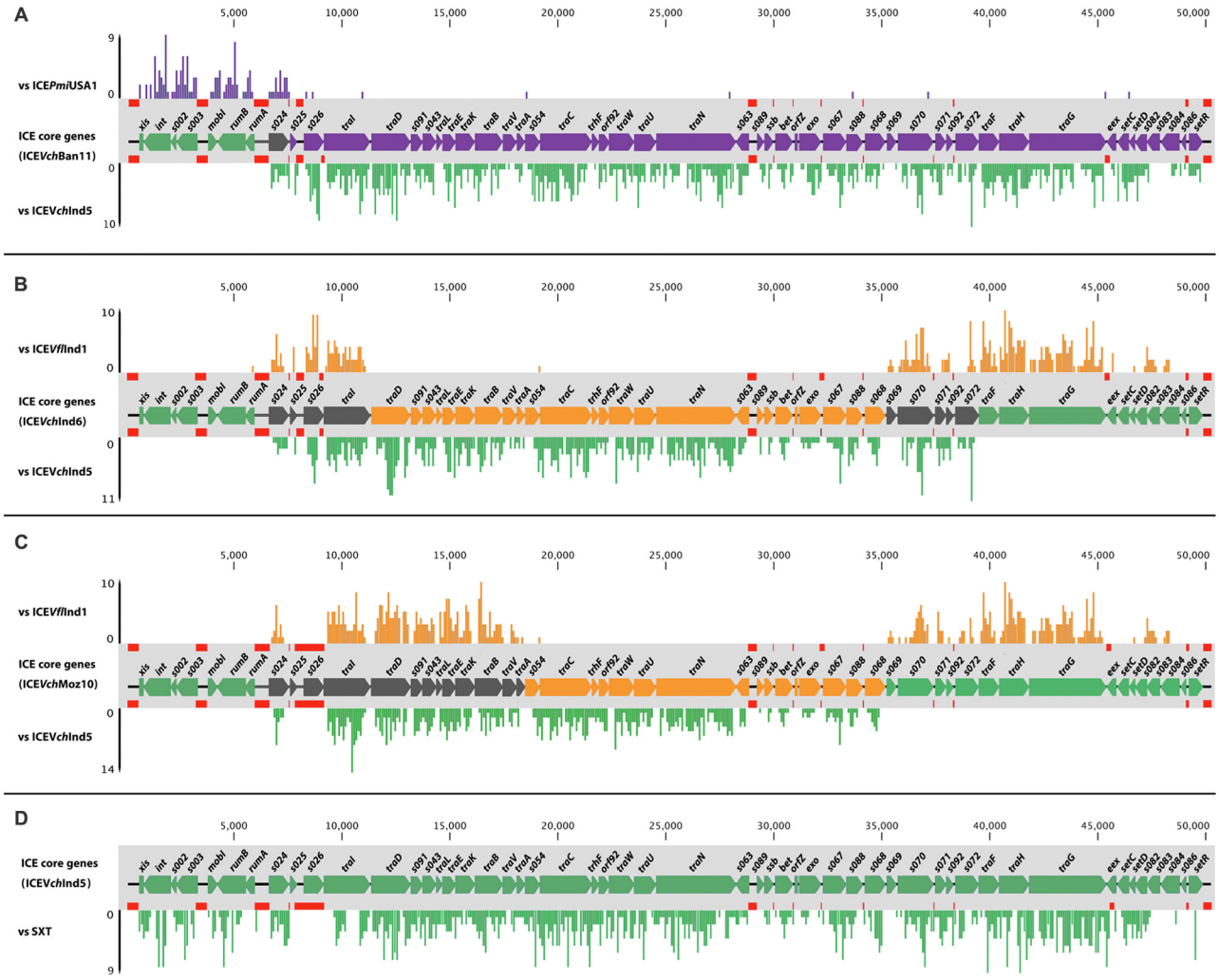


**FIG 3** Signatures of recombination between ICEs in *V. cholerae*. The distribution of SNPs is shown for each ICE profile with respect to the 52 core genes of different reference ICEs. The *y* axis data give the number of SNPs within a 100-bp window. (A) ICE*Vch*Ban11 reads (accession number ERS016137) are mapped against ICE*Pmi*USA1 (purple) and ICE*Vch*Ind5 (green). (B) ICE*Vch*Ind6 reads (accession number ERS013257) are mapped against ICE*Vfl*Ind1 (orange) and ICE*Vch*Ind5 (green). (C) ICE*Vch*Moz10 reads (accession number ERS013126) are mapped against ICE*Vfl*Ind1 (orange) and ICE*Vch*Ind5 (green). (D) SXT reads (accession number ERS013124) are mapped against ICE*Vch*Ind5. Genes with high (>98%) sequence similarity to the respective reference ICE are represented in the same color. No recent recombination signatures were detected in dark-gray genes. Red bars depict low coverage and intergenic regions where SNPs could not be called (<5× coverage, >25 bp).

In this paper, we demonstrate that both multiple acquisitions and homologous recombination were important in shaping the current ICE diversity in the *V. cholerae* seventh-pandemic lineage. We infer that two distinct elements, ICE*Vch*Ind5 and SXT, were independently acquired by ICE-free *V. cholerae*. After that, homologous recombination with other ICEs, likely acquired by conjugation, played an important role in further shaping the ICE genetic architecture and promoting the formation of new hybrid elements in *V. cholerae* O1 El Tor.

We classified ICEs currently circulating among *V. cholerae* clinical strains into at least five different groups, representing elements with identical genetic cargoes in their variable regions (Fig. 2). Genetic variation within ICE groups is low, with the exception of a few major deletions, mostly located in the antibiotic resistance cluster (Fig. 1). Similar deletions were documented in ICEs from the Haitian epidemic from 2011 and 2012, indicating that this region is prone to frequent rearrangements (37). This is not entirely surprising given the highly recombinogenic nature of the ICE antibiotic resistance cluster encoding Tn*3*-like and ISCR*2* elements recognized as powerful gene capture and movement tools (26, 38, 39).

The phylogenetic methods we employed to test for single versus multiple ICE acquisitions and to estimate the dates of these events rely on the analyzed sequences not having been affected by recent genetic recombination. Previous studies have demonstrated that SXT-R391 ICEs generate their own genetic diversity by mediating recombination between two ICEs arranged in a tandem array (23). ICE-encoded proteins Bet and Exo, working in a RecA-independent homologous recombination system, mediate the formation of hybrid ICEs and are probably responsible for most of the genetic variability observed in the SXT-R391 family (22). However, the operation of this process with identical or sim-

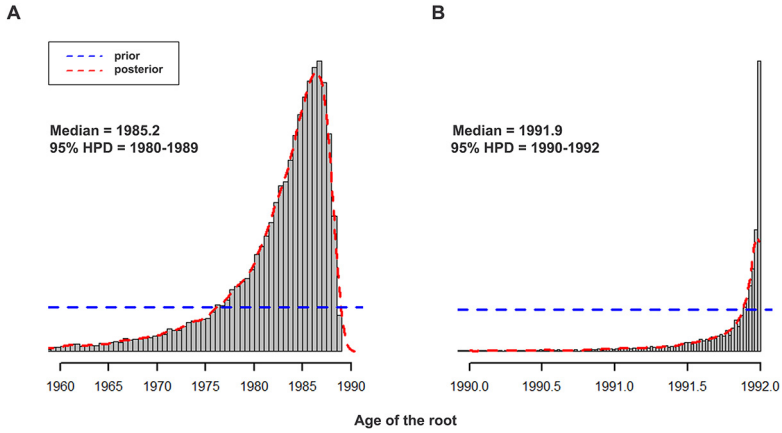


**FIG 4** Estimated date for the acquisition of ICE*Vch*Ind5 and SXT in *V. cholerae*. Data represent the posterior-probability distributions (red dashed lines) for the time to most recent common ancestor (TMRCA) of ICE*Vch*Ind5 (A) and SXT (B). Median values and 95% confidence intervals are reported in the legend beside the graph. The *x* axis shows the age of the root in years. Flat priors (blue dashed line) were applied in BEAST, as the tree was calibrated using tip dates only (see Materials and Methods for details).

ilar elements is inhibited by the ICE entry exclusion system, which prevents redundant transfer by converting the host cells into poor recipients (25). As a consequence of this constraint, homologous recombination is allowed only between ICEs of different exclusion groups. Moreover, ICE loss from the host chromosome followed by reacquisition of another ICE is a most unlikely event as all the ICEs described in *V. cholerae* encode a toxin-antitoxin system that ensures ICE maintenance (40).

Data indicating the absence of substantial recombination within exclusion groups and secondary loss of ICEs are supported by the temporal patterns we report for distribution of ICEs and their genetic organization in the seventh-pandemic lineage. While we detected extensive recombination post-ICE acquisition by seventh-pandemic *V. cholerae* strains, the backbone genes of ICE*Vch*Ind5 and SXT did not show any recent signature of inter-ICE homologous recombination (Fig. 3D). Our analyses of ICE alignments dated the acquisitions of ICE*Vch*Ind5 and SXT to 1985 and 1992, respectively. Comparison with the *V. cholerae* genome phylogeny proposed by Mutreja et al. (8) indicates that the first isolates carrying ICE*Vch*Ind5 are represented by wave 2 strains, an outgroup of the wave 3 lineage, where ICE*Vch*Ind5 is dominant. The MRCA of 1981 for O1 and O139 serogroups used as a proxy for the acquisition of the SXT-R391 ICEs by Mutreja et al. (8) predates our estimation for the acquisition of ICE*Vch*Ind5 (1985), although the confidence intervals partially overlap. However, the date proposed by Mutreja et al. does not correspond to the acquisition of SXT and/or the O139 antigen but marks only the split between the El Tor progenitor of the O139 serogroup and the rest of the seventh-pandemic lineage. Interestingly, when we looked at the two O1 strains closest to the O139 sublineage (strains PRL5 and A109 [accession no. ERS013145 and ERS013173, respectively]) (8), we found no ICEs in their genomes. Our estimation for the acquisition date of SXT is set 10 years later and precedes the first outbreak of *V. cholerae* O139 in the Indian subcontinent by just 1 year. While we are confident about the accuracy of the dating of ICE*Vch*Ind5 acquisition, our date for SXT should be treated with caution because it was estimated using the only four SXT element sequences currently available. This small set may not capture the entirety of the genetic diversity of the clade and hence may correspond to a more recent MRCA. Thus, the date of acquisition for SXT presented here is likely to represent an underestimation.

In addition to SXT and ICE*Vch*Ind5, we detected significant recent inter-ICE homologous recombination in our analysis of other ICE profiles. Network analysis (see Fig. S3 in the supplemental material) and analysis of the distribution of SNPs along 52 conserved ICE core genes (Fig. 3) revealed that ICE*Vch*Ind6, ICE*Vch*Ban11, and ICE*Vch*Moz10 are likely the result of recent inter-ICE homologous recombination events involving ICE*Vch*Ind5 and other ICEs circulating in other *Gammaproteobacteria*, such as ICE*Pmi*USA1 and ICE*Vfl*Ind1 (Fig. 3) isolated from *P. mirabilis* and *V. fluvialis*, respectively. While it is difficult to identify recombination donors unequivocally because of the mosaic nature of SXT-R391 ICEs, the observed SNP patterns strongly support the hypothesis that the three ICEs were recently generated within *V. cholerae* strains likely previously carrying ICE*Vch*Ind5 (Fig. 3). No doubt, the aquatic environment and its bacterial inhabitants provide SXT-R391 ICEs with a genetic pool from which to draw new genetic material and expand their gene repertoire.

Our data indicate that ICE acquisition by seventh-pandemic *V. cholerae* was not a single event but that SXT and ICE*Vch*Ind5 were independently acquired by ICE-free *V. cholerae* strains. A schematic representation of the acquisition of the SXT-R391 ICEs is presented in Fig. 5. Furthermore, the analyses presented here reveal that ICE acquisition by *V. cholerae* promoted subsequent ICE rearrangements and generated significant variability on the genomic scale. In fact, the ability of SXT-R391 ICEs to mobilize large regions of host DNA and/or ICE-associated genomic islands was previously documented (41). Katz et al. (37) reported a large (~400-kb) inversion around the ICE integration region of a Haitian *V. cholerae* isolate. The recent description of ICE-encoded extended-spectrum cephalosporin resistance ($bla_{CMY-2}$) in a clinical *Proteus mirabilis* strain isolated in Japan (42) further highlights the ability of this family of mobile genetic elements to acquire and spread novel drug resistance genes. The impact of ICE acquisition gains noteworthy relevance in light of recent findings showing *V. cholerae* seventh-pandemic clones to be poorly transformable (37).

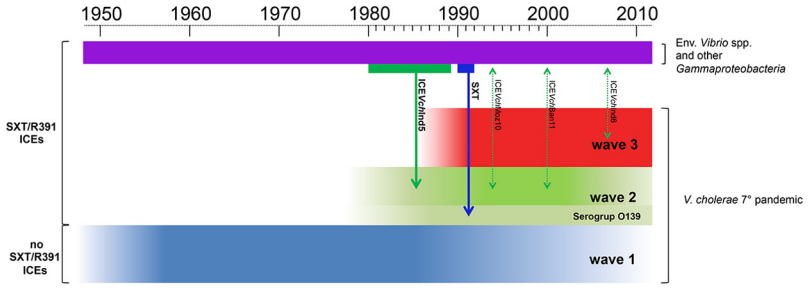


**FIG 5** Schematic representation of the acquisition of SXT-R391 ICEs in seventh-pandemic *V. cholerae*. The three overlapping pandemic waves of global transmission are depicted in blue (first wave), green (second wave), and red (third wave), according to the color scheme and the inferred dates from Mutreja et al. (8). The O139 serogroup is shown as a subset of the second pandemic wave, in light green. Vertical arrows depict the acquisition events from the environment and subsequent rearrangements by recombination with other SXT-R391 ICEs acquired by conjugation (hypothetical bacterial donors are depicted using a purple bar). The likely acquisition dates of ICE*Vch*Ind5 and SXT are shown by green and blue arrows, respectively, with color bars spanning the 95% confidence intervals estimated by our analysis of the ICE sequences (see text and Fig. 4B). Subsequent inter-ICE recombination events giving rise to new hybrid ICEs are depicted as green dashed double arrows, with corresponding dates associated with the oldest isolate showing the recombinant ICE. Env., environmental.

In summary, these results illustrate how ICE integration promotes both element-associated and chromosomal rearrangement, thus acting as a horizontal gene transfer hot spot for the host strain.

## MATERIALS AND METHODS

**ICE data set creation, *de novo* assembly, and comparative analysis.** The NCBI archive and the Sequence Read Archive (SRA) were screened for SXT-R391 ICE sequences. A list of all genome sequences and reference ICEs used in this study is presented in Table S1. All downloaded reads were refiltered and trimmed using a FASTX-Toolkit with a minimum Q20 quality-score cutoff before performing the assembly step. The resultant data set was then assembled *de novo* using Velvet 1/1/06 (43). Contigs related to ICE*Vch*Ind6 and ICE*Vch*Ban11 were extracted from the genomic assemblies of *V. cholerae* 4605 and *V. cholerae* 4672 (accession no. ERS013257 and ERS016137, respectively) and used as references for further analysis. The backbone sequences of ICE*Vch*Ind5, SXT, R391, ICE*Pda*Spa1, ICE*Vch*Hai2, ICE*Vch*Mex1, ICE*Spu*PO1, ICE*Pal*Ban1, ICE*Pmi*USA1, ICE*Vch*Ban11, ICE*Vfl*Ind1, ICE*Vch*Ind6, and ICE*Vch*Moz10 used for recombination analysis were built by concatenating ICE core regions and removing all variable genes in Artemis (44). BLAST atlas maps were built by comparing concatenated *de novo* assemblies with reference sequences of ICE*Vch*Ind5 (GQ463142), ICE*Vch*Moz10 (ACHZ00000000), and SXT (AY055428) using BLASTn with a >90% identity threshold within the CGView Comparison Tool (30, 45). The Artemis Comparison Tool and CGView were used for further comparisons and to visualize the maps (45–47).

**Variant calling and homologous recombination analysis.** Reads were organized in groups reflecting similarity with their reference sequences based on the results of the analysis performed on the *de novo* assembly. Bowtie2 (48) was used to map processed reads belonging to samples of each ICE group against their reference sequences (ICE*Vch*Ind5, ICE*Vch*Ind6, ICE*Vch*Moz10, and SXT) and all 96 samples against ICE*Vch*Ind5. The resultant sam files were converted to bam, coordinate sorted with Picard Tools (http://picard.sourceforge.net), and imported in CLC Genomics Workbench 6.0 for variant calling. CLC Probabilistic Variant Caller was used for this purpose with a variant probability threshold of 90% and with minimum sequence coverage of 10×. SNP calling for recombination analysis was performed with the same settings described above, mapping representative reads corresponding to ICE*Vch*Ban11 (accession number ERS016137), ICE*Vch*Ind6 (accession number ERS013257), ICE*Vch*Moz10 (accession number ERS013126), and SXT (accession number ERS013124) profiles against the 52 core genes of ICE*Vch*Ind5 and a set of eight additional ICEs found in non-O1/non-O139 *V. cholerae* strains and other *Gammaproteobacteria* (see Table S1 in the supplemental material). SNPs were filtered to include only synonymous SNPs to avoid potential selection bias introduced by strong selective pressure on specific ICE backbone genes. SNPs were called only on coding regions to avoid length discrepancies in the intergenic regions of the ICE backbone.

The initial search for recombination was conducted by building neighbor-net and split-decomposition networks using the SplitsTree4 software (33). A Phi test for recombination was applied using a permutation test of 100,000 iterations and PhiPack software (49). Additional tests for recombination detection were conducted using Recombination Detection Program 4 (RDP4) v.4.35 (36) and five different methods (RDP, MaxChi, Bootscan, Chimaera, and SiScan). All the tests were executed with default parameters, and the multiple-comparison corrected *P*-value cutoff value was set to 0.05.

**Bayesian analysis.** ICE consensus sequences for each sample were extracted from the reads mapped against ICE*Vch*Ind5 and SXT using a coverage threshold of >10× and the base-calling quality scores for handling SNP conflicts as implemented by CLC Genomic Workbench. Sequences were then aligned with MUSCLE (50). PartitionFinder (51) was used to estimate the best-fit nucleotide substitution model for each nucleotide subset. Bayesian phylogenetic analyses were performed with BEAST 1.7.4 (35) on both whole elements and various nucleotide partitions (coding positions 1 and 2, coding position 3, noncoding sites, and pseudogenes). Substitution and clock models were unlinked in BEAST, while the tree topologies of the four nucleotide subsets were assumed to be the same. A separate substitution model and molecular clock but a single tree was estimated for each partition. We assumed a lognormal relaxed clock to allow variation in rates among branches. To minimize prior assumptions about demographic history, we adopted a Bayesian skyline plot approach in order to integrate data over different coalescent histories. Rate variations among sites were modeled with a discrete gamma distribution with four rate categories. The tree was calibrated using tip dates only, with sample time spans ranging from 1992 to 2002 and from 1989 to 2010 for SXT-like and ICE*Vch*Ind5-like elements, respectively. BEAST analyses were performed on two datasets. The first contained 71 ICE*Vch*Ind5-like element sequences and the second the 4 SXT-like sequences. A third run was performed on a subset of 62 ICEs of the ICE*Vch*Ind5 group to validate the results. Flat priors (i.e., uniform distributions) were applied for the substitution rate ($1.10^{-10}$ to $1.10^{-3}$ substitutions/site/year) as well as for the age of any node in the tree, including the height of the root (1500 to 1989 and 1500 to 1992 for the MRCAs of the ICE*Vch*Ind5 and SXT groups, respectively). Posterior distributions of parameters, including divergence times and substitution rates, were estimated by Markov chain Monte Carlo (MCMC) sampling in BEAST. For each analysis that we ran, we combined five independent chains from which samples were drawn

every 5,000 MCMC steps from a total of 50,000,000 steps, after a burn-in count of 5,000,000 steps was discarded. Convergence to the stationary distribution and sufficient sampling were checked by inspection of posterior samples. The best-supported tree was estimated from the combined samples using the Maximum clade credibility method implemented in Tree Annotator after a burn-in level of 10% was discarded.

## SUPPLEMENTAL MATERIAL

Supplemental material for this article may be found at http://mbio.asm.org/lookup/suppl/doi:10.1128/mBio.01356-14/-/DCSupplemental.

Figure S1, TIF file, 1.1 MB.
Figure S2, TIF file, 0.2 MB.
Figure S3, TIF file, 0.8 MB.
Figure S4, TIF file, 0.2 MB.
Table S1, PDF file, 0.3 MB.

## ACKNOWLEDGMENTS

This work was supported by European Research Council grant ERC 260801—BIG_IDEA (to F.B.). M.S. was supported by a fellowship from the Institute Pasteur—Fondazione Cenci Bolognetti, Rome, Italy. D.C. was supported by National Institutes of Health grant no. 2RO1A1039129-11A2.

We are thankful to Genevieve Garriss (Institute Pasteur, Paris, France) for insightful comments on the manuscript.

## REFERENCES

1. **Kaper JB, Morris JG, Levine MM.** 1995. Cholera. Clin. Microbiol. Rev. **8:**48–86.
2. **Harris JB, LaRocque RC, Qadri F, Ryan ET, Calderwood SB.** 2012. Cholera. Lancet **379:**2466–2476. http://dx.doi.org/10.1016/S0140-6736(12)60436-X.
3. **Faruque SM, Sack DA, Sack RB, Colwell RR, Takeda Y, Nair GB.** 2003. Emergence and evolution of *Vibrio cholerae* O139. Proc. Natl. Acad. Sci. U. S. A. **100:**1304–1309. http://dx.doi.org/10.1073/pnas.0337468100.
4. **Blokesch M, Schoolnik GK.** 2007. Serogroup conversion of *Vibrio cholerae* in aquatic reservoirs. PLoS Pathog. **3:**e81. http://dx.doi.org/10.1371/journal.ppat.0030081.
5. **Chun J, Grim CJ, Hasan NA, Lee JH, Choi SY, Haley BJ, Taviani E, Jeon YS, Kim DW, Lee JH, Brettin TS, Bruce DC, Challacombe JF, Detter JC, Han CS, Munk AC, Chertkov O, Meincke L, Saunders E, Walters RA, Huq A, Nair GB, Colwell RR.** 2009. Comparative genomics reveals mechanism for short-term and long-term clonal transitions in pandemic *Vibrio cholerae*. Proc. Natl. Acad. Sci. U. S. A. **106:**15442–15447. http://dx.doi.org/10.1073/pnas.0907787106.
6. **Safa A, Nair GB, Kong RY.** 2010. Evolution of new variants of *Vibrio cholerae* O1. Trends Microbiol. **18:**46–54. http://dx.doi.org/10.1016/j.tim.2009.10.003.
7. **Reimer AR, Van Domselaar G, Stroika S, Walker M, Kent H, Tarr C, Talkington D, Rowe L, Olsen-Rasmussen M, Frace M, Sammons S, Dahourou GA, Boncy J, Smith AM, Mabon P, Petkau A, Graham M, Gilmour MW, Gerner-Smidt P, V. cholerae Outbreak Genomics Task Force.** 2011. Comparative genomics of *Vibrio cholerae* from Haiti, Asia, and Africa. Emerg. Infect. Dis. **17:**2113–2121. http://dx.doi.org/10.3201/eid1711.110794.
8. **Mutreja A, Kim DW, Thomson NR, Connor TR, Lee JH, Kariuki S, Croucher NJ, Choi SY, Harris SR, Lebens M, Niyogi SK, Kim EJ, Ramamurthy T, Chun J, Wood JL, Clemens JD, Czerkinsky C, Nair GB, Holmgren J, Parkhill J, Dougan G.** 2011. Evidence for several waves of global transmission in the seventh cholera pandemic. Nature **477:**462–465. http://dx.doi.org/10.1038/nature10392.
9. **Cho YJ, Yi H, Lee JH, Kim DW, Chun J.** 2010. Genomic evolution of *Vibrio cholerae*. Curr. Opin. Microbiol. **13:**646–651. http://dx.doi.org/10.1016/j.mib.2010.08.007.
10. **Wozniak RA, Fouts DE, Spagnoletti M, Colombo MM, Ceccarelli D, Garriss G, Déry C, Burrus V, Waldor MK.** 2009. Comparative ICE genomics: insights into the evolution of the SXT/R391 family of ICEs. PLoS Genet. **5:**e1000786. http://dx.doi.org/10.1371/journal.pgen.1000786.
11. **Badhai J, Kumari P, Krishnan P, Ramamurthy T, Das SK.** 2013. Presence of SXT integrating conjugative element in marine bacteria isolated from the mucus of the coral *Fungia echinata* from Andaman Sea. FEMS Microbiol. Lett. **338:**118–123. http://dx.doi.org/10.1111/1574-6968.12033.
12. **Rodríguez-Blanco A, Lemos ML, Osorio CR.** 2012. Integrating conjugative elements as vectors of antibiotic, mercury, and quaternary ammonium compound resistance in marine aquaculture environments. Antimicrob. Agents Chemother. **56:**2619–2626. http://dx.doi.org/10.1128/AAC.05997-11.
13. **Waldor MK, Tschäpe H, Mekalanos JJ.** 1996. A new type of conjugative transposon encodes resistance to sulfamethoxazole, trimethoprim, and streptomycin in *Vibrio cholerae* O139. J. Bacteriol. **178:**4157–4165.
14. **Peters SE, Hobman JL, Strike P, Ritchie DA.** 1991. Novel mercury resistance determinants carried by IncJ plasmids pMERPH and R391. Mol. Gen. Genet. **228:**294–299.
15. **Wozniak RA, Waldor MK.** 2010. Integrative and conjugative elements: mosaic mobile genetic elements enabling dynamic lateral gene flow. Nat. Rev. Microbiol. **8:**552–563. http://dx.doi.org/10.1038/nrmicro2382.
16. **Pembroke JT, Piterina AV.** 2006. A novel ICE in the genome of *Shewanella putrefaciens* W3-18-1: comparison with the SXT/R391 ICE-like elements. FEMS Microbiol. Lett. **264:**80–88. http://dx.doi.org/10.1111/j.1574-6968.2006.00452.x.
17. **Bordeleau E, Brouillette E, Robichaud N, Burrus V.** 2010. Beyond antibiotic resistance: integrating conjugative elements of the SXT/R391 family that encode novel diguanylate cyclases participate to c-di-GMP signalling in *Vibrio cholerae*. Environ. Microbiol. **12:**510–523. http://dx.doi.org/10.1111/j.1462-2920.2009.02094.x.
18. **Burrus V, Quezada-Calvillo R, Marrero J, Waldor MK.** 2006. SXT-related integrating conjugative element in New World *Vibrio cholerae*. Appl. Environ. Microbiol. **72:**3054–3057. http://dx.doi.org/10.1128/AEM.72.4.3054-3057.2006.
19. **Ceccarelli D, Spagnoletti M, Bacciu D, Danin-Poleg Y, Mendiratta DK, Kashi Y, Cappuccinelli P, Burrus V, Colombo MM.** 2011. ICE*Vch*Ind5 is prevalent in epidemic *Vibrio cholerae* O1 El Tor strains isolated in India. Int. J. Med. Microbiol. **301:**318–324. http://dx.doi.org/10.1016/j.ijmm.2010.11.005.
20. **Ceccarelli D, Spagnoletti M, Bacciu D, Cappuccinelli P, Colombo MM.** 2011. New *V. cholerae* atypical el Tor variant emerged during the 2006 epidemic outbreak in Angola. BMC Microbiol. **11:**130. http://dx.doi.org/10.1186/1471-2180-11-130.
21. **Ceccarelli D, Spagnoletti M, Cappuccinelli P, Burrus V, Colombo MM.** 2011. Origin of *Vibrio cholerae* in Haiti. Lancet Infect. Dis. **11:**262. http://dx.doi.org/10.1016/S1473-3099(11)70079-2.
22. **Garriss G, Waldor MK, Burrus V.** 2009. Mobile antibiotic resistance encoding elements promote their own diversity. PLoS Genet. **5:**e1000775. http://dx.doi.org/10.1371/journal.pgen.1000775.
23. **Burrus V, Waldor MK.** 2004. Formation of SXT tandem arrays and SXT-R391 hybrids. J. Bacteriol. **186:**2636–2645. http://dx.doi.org/10.1128/JB.186.9.2636-2645.2004.
24. **Garriss G, Poulin-Laprade D, Burrus V.** 2013. DNA-damaging agents induce the RecA-independent homologous recombination functions of integrating conjugative elements of the SXT/R391 family. J. Bacteriol. **195:**1991–2003. http://dx.doi.org/10.1128/JB.02090-12.
25. **Marrero J, Waldor MK.** 2007. The SXT/R391 family of integrative conjugative elements is composed of two exclusion groups. J. Bacteriol. **189:**3302–3305. http://dx.doi.org/10.1128/JB.01902-06.
26. **Toleman MA, Bennett PM, Walsh TR.** 2006. ISCR elements: novel gene-capturing systems of the 21st century? Microbiol. Mol. Biol. Rev. **70:**296–316. http://dx.doi.org/10.1128/MMBR.00048-05.
27. **Redfield RJ.** 2001. Do bacteria have sex? Nat. Rev. Genet. **2:**634–639. http://dx.doi.org/10.1038/35084593.
28. **Beaber JW, Hochhut B, Waldor MK.** 2002. Genomic and functional analyses of SXT, an integrating antibiotic resistance gene transfer element derived from *Vibrio cholerae*. J. Bacteriol. **184:**4259–4269. http://dx.doi.org/10.1128/JB.184.15.4259-4269.2002.
29. **McLeod SM, Burrus V, Waldor MK.** 2006. Requirement for *Vibrio cholerae* integration host factor in conjugative DNA transfer. J. Bacteriol. **188:**5704–5711. http://dx.doi.org/10.1128/JB.00564-06.
30. **Camacho C, Coulouris G, Avagyan V, Ma N, Papadopoulos J, Bealer K, Madden TL.** 2009. BLAST+: architecture and applications. BMC Bioinformatics **10:**421. http://dx.doi.org/10.1186/1471-2105-10-421.
31. **Hendriksen RS, Price LB, Schupp JM, Gillece JD, Kaas RS, Engelthaler DM, Bortolaia V, Pearson T, Waters AE, Upadhyay BP, Shrestha SD, Adhikari S, Shakya G, Keim PS, Aarestrup FM.** 2011. Population genet-

ics of *Vibrio cholerae* from Nepal in 2010: evidence on the origin of the Haitian outbreak. mBio 2(4):e00157–11. http://dx.doi.org/10.1128/mBio.00157-11.
32. **Daccord A, Ceccarelli D, Rodrigue S, Burrus V.** 2012. Comparative analysis of mobilizable genomic islands. J. Bacteriol. **195:**1–10. http://dx.doi.org/10.1128/JB.01985-12.PubMed.
33. **Huson DH, Bryant D.** 2006. Application of phylogenetic networks in evolutionary studies. Mol. Biol. Evol. **23:**254–267. http://dx.doi.org/10.1093/molbev/msj030.PubMed.
34. **He M, Sebaihia M, Lawley TD, Stabler RA, Dawson LF, Martin MJ, Holt KE, Seth-Smith HM, Quail MA, Rance R, Brooks K, Churcher C, Harris D, Bentley SD, Burrows C, Clark L, Corton C, Murray V, Rose G, Thurston S, van Tonder A, Walker D, Wren BW, Dougan G, Parkhill J.** 2010. Evolutionary dynamics of *Clostridium difficile* over short and long time scales. Proc. Natl. Acad. Sci. U. S. A. **107:**7527–7532. http://dx.doi.org/10.1073/pnas.0914322107.
35. **Drummond AJ, Rambaut A.** 2007. BEAST: Bayesian evolutionary analysis by sampling trees. BMC Evol. Biol. **7:**214. http://dx.doi.org/10.1186/1471-2148-7-214.
36. **Martin DP, Lemey P, Lott M, Moulton V, Posada D, Lefeuvre P.** 2010. RDP3: a flexible and fast computer program for analyzing recombination. Bioinformatics **26:**2462–2463. http://dx.doi.org/10.1093/bioinformatics/btq467.
37. **Katz LS, Petkau A, Beaulaurier J, Tyler S, Antonova ES, Turnsek MA, Guo Y, Wang S, Paxinos EE, Orata F, Gladney LM, Stroika S, Folster JP, Rowe L, Freeman MM, Knox N, Frace M, Boncy J, Graham M, Hammer BK, Boucher Y, Bashir A, Hanage WP, Van Domselaar G, Tarr L.** 2013. Evolutionary dynamics of *Vibrio cholerae* O1 following a single-source introduction to Haiti. mBio **4:**e00398-10. doi: http://dx.doi.org/10.1128/mBio.00398-13.
38. **Toleman MA, Walsh TR.** 2011. Combinatorial events of insertion sequences and ICE in gram-negative bacteria. FEMS Microbiol. Rev. **35:**912–935. http://dx.doi.org/10.1111/j.1574-6976.2011.00294.x.
39. **Toleman MA, Walsh TR.** 2010. ISCR elements are key players in IncA/C plasmid evolution. Antimicrob. Agents Chemother. **54:**; author reply, 3534. http://dx.doi.org/10.1128/AAC.00383-103534.
40. **Wozniak RA, Waldor MK.** 2009. A toxin-antitoxin system promotes the maintenance of an integrative conjugative element. PLoS Genet. **5:**e1000439. http://dx.doi.org/10.1371/journal.pgen.1000439.
41. **Daccord A, Ceccarelli D, Burrus V.** 2010. Integrating conjugative elements of the SXT/R391 family trigger the excision and drive the mobilization of a new class of *Vibrio* genomic islands. Mol. Microbiol. **78:**576–588. http://dx.doi.org/10.1111/j.1365-2958.2010.07364.x.
42. **Harada S, Ishii Y, Saga T, Tateda K, Yamaguchi K.** 2010. Chromosomally encoded blaCMY-2 located on a novel SXT/R391-related integrating conjugative element in a Proteus mirabilis clinical isolate. Antimicrob. Agents Chemother. **54:**3545–3550. http://dx.doi.org/10.1128/AAC.00111-10.
43. **Zerbino DR, Birney E.** 2008. Velvet: algorithms for de novo short read assembly using de Bruijn graphs. Genome Res. **18:**821–829. http://dx.doi.org/10.1101/gr.074492.107.
44. **Rutherford K, Parkhill J, Crook J, Horsnell T, Rice P, Rajandream MA, Barrell B.** 2000. Artemis: sequence visualization and annotation. Bioinformatics **16:**944–945. http://dx.doi.org/10.1093/bioinformatics/16.10.944.
45. **Stothard P, Grant JR, Arantes AS.** 2012. Comparing thousands of circular genomes using the CGView Comparison Tool. BMC Genomics **13:**202. http://dx.doi.org/10.1186/1471-2164-13-202.PubMed.
46. **Stothard P, Wishart DS.** 2005. Circular genome visualization and exploration using CGView. Bioinformatics **21:**537–539. http://dx.doi.org/10.1093/bioinformatics/bti054.
47. **Carver TJ, Rutherford KM, Berriman M, Rajandream MA, Barrell BG, Parkhill J.** 2005. ACT: the Artemis Comparison Tool. Bioinformatics **21:**3422–3423. http://dx.doi.org/10.1093/bioinformatics/bti553.
48. **Langmead B, Salzberg SL.** 2012. Fast gapped-read alignment with Bowtie 2. Nat. Methods **9:**357–359. http://dx.doi.org/10.1038/nmeth.1923.
49. **Bruen TC, Philippe H, Bryant D.** 2006. A simple and robust statistical test for detecting the presence of recombination. Genetics **172:**2665–2681. http://dx.doi.org/10.1534/genetics.105.048975.PubMed.
50. **Edgar RC.** 2004. MUSCLE: multiple sequence alignment with high accuracy and high throughput. Nucleic Acids Res. **32:**1792–1797. http://dx.doi.org/10.1093/nar/gkh340.
51. **Lanfear R, Calcott B, Ho SY, Guindon S.** 2012. Partitionfinder: combined selection of partitioning schemes and substitution models for phylogenetic analyses. Mol. Biol. Evol. **29:**1695–1701. http://dx.doi.org/10.1093/molbev/mss020.